The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

The effects of ageing and visual noise on conceptual integration during sentence reading
Xuefei Gao a, Brian R. Levinthal b & Elizabeth A. L. Stine-Morrow a
a Beckman Institute & Department of Educational Psychology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
b Department of Psychology, Northwestern University, Evanston, IL, USA
Accepted author version posted online: 15 Mar 2012. Published online: 25 Apr 2012.


To link to this article: http://dx.doi.org/10.1080/17470218.2012.674146

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
The effects of ageing and visual noise on conceptual integration during sentence reading

Xuefei Gao¹, Brian R. Levinthal², and Elizabeth A. L. Stine-Morrow¹

¹Beckman Institute & Department of Educational Psychology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
²Department of Psychology, Northwestern University, Evanston, IL, USA

The effortfulness hypothesis implies that difficulty in decoding the surface form, as in the case of age-related sensory limitations or background noise, consumes the attentional resources that are then unavailable for semantic integration in language comprehension. Because ageing is associated with sensory declines, degrading of the surface form by a noisy background can pose an extra challenge for older adults. In two experiments, this hypothesis was tested in a self-paced moving window paradigm in which younger and older readers' online allocation of attentional resources to surface decoding and semantic integration was measured as they read sentences embedded in varying levels of visual noise. When visual noise was moderate (Experiment 1), resource allocation among young adults was unaffected but older adults allocated more resources to decode the surface form at the cost of resources that would otherwise be available for semantic processing; when visual noise was relatively intense (Experiment 2), both younger and older participants allocated more attention to the surface form and less attention to semantic processing. The decrease in attentional allocation to semantic integration resulted in reduced recall of core ideas in both experiments, suggesting that a less organized semantic representation was constructed in noise. The greater vulnerability of older adults at relatively low levels of noise is consistent with the effortfulness hypothesis.

Keywords: Ageing; Visual noise; Conceptual integration; Reading.

Ageing is associated with sensory declines, which lead to difficulty in identifying and discriminating among visual and auditory stimuli (Pichora-Fuller, Schneider, & Daneman, 1995; Speranza, Daneman, & Schneider, 2000). According to the effortfulness hypothesis (Rabbitt, 1968; Wingfield, Tun, & McCoy, 2005), sensory challenges created by a muddy signal or by ageing could make perceptual processes effortful and consume attentional resources that would otherwise...

Correspondence should be addressed to Xuefei Gao, Beckman Institute, University of Illinois, Urbana, IL USA. E-mail: gao5@illinois.edu

This study was supported by the National Institute on Aging (Grant R01 AG13935) to whom we are most grateful. We also would like to thank Kara Federmeier, Kiel Christianson, Soo Rim Noh, and Matthew Shake for insightful discussions and comments on earlier drafts of this paper. Appreciation also goes to Micaela Chan and Minerva Dorant for testing the participants and transcribing the auditory files.

© 2012 The Experimental Psychology Society

be used for higher cognitive functions, such as the semantic analysis of language. The current studies were designed to directly test this hypothesis in reading by embedding text in visual noise (Speranza et al., 2000) and measuring semantic processing using a resource allocation paradigm (Stine-Morrow, Milinder, Pullara, & Herman, 2001; Stine-Morrow & Miller, 2009; Stine-Morrow, Miller, Gagne, & Hertzog, 2008; Stine-Morrow, Miller, & Hertzog, 2006).

**Effortfulness hypothesis**

The effortfulness hypothesis was first proposed by Rabbitt (1968), who tested young adults with normal hearing in a digit recall task. Participants were presented with two lists of digits auditorily. The first list was always presented in a quiet background (without masking), and the second one was heard either in a quiet or in a noisy background. He found that participants’ recall of the first list was impaired if it was followed by a list that was embedded in noise. Thus, he concluded that the impaired memory for digits was due to the deprivation of processing resources for rehearsing the first list, which was caused by the decoding difficulty of the second list embedded in noise.

Wingfield and his colleagues (2005; McCoy et al., 2005) extended the effortfulness hypothesis to the effects of hearing loss among older adults. They reported that when two groups of elders, one with relatively good hearing and the other with mild-to-moderate hearing loss, were required to recall an auditorily presented three-word list, both groups performed equally well at recalling the final word of the list, implying that both groups were capable of extracting and identifying the auditory input. However, the recall of the first two words in the list was differentially depressed among the hearing impaired (see also Rabbitt, 1991). Those counterintuitive findings, together with other empirical evidence (Heinrich, Schneider, & Craik, 2008; Murphy, Craik, Li, & Schneider, 2000), obtained by using digit or word lists as experimental materials, provided support for the effortfulness hypothesis—namely, that sensory challenge (i.e., muddy signal or hearing loss) can tax resources for higher order cognition (i.e., organization and elaboration processes that support memory).

There are two types of “noise” that can impact memory performance: internal noise, which is created by ageing, disease, or brain damage (Allen, 1990); and external noise, created by environmental conditions that degrade sensory input, such as a blurry screen presentation or speech in a babble background. Sensory decline is only one sort of internal noise resulting from advancing age. In a memory-scanning task, Allen (1990) presented younger and older adults with strings of letters in which the distance between serial positions of the letter at encoding and retrieval was varied. The task was to verify at retrieval whether a letter appeared in the same location as it did during encoding. He found that for both younger and older adults, reaction time and error rate increased as the distance between letter positions decreased from encoding to retrieval. Because older adults showed this effect more strongly, he concluded that older adults might experience greater “processing variability”, which he characterized as “internal noise” or “neural noise”, in episodic memory. Notably, internal and external noise appear to produce analogous effects: The hearing-impaired older adults in Wingfield et al.’s study (2005) behaved in a way that was quite similar to the young adults in the noisy condition of Rabbitt’s (1968) study, with regard to memory performance. Together with other findings (Brown & Pichora-Fuller, 2000; Murphy et al., 2000; Rayner, Reichele, Stroud, Williams, & Pollatsek, 2006; Schneider, Daneman, Murphy, & Kwong See, 2000), these results support the idea that sensory declines and external noise can similarly impact performance by shifting the processing resources that are needed for information processing in working memory to effort at decoding of the surface form.

**The effortfulness hypothesis in sentence processing**

Several researchers have attempted to test the effortfulness hypothesis in language processing. Using a simulated visual impairment technique
with normally sighted younger adults, Dickinson and Rabbitt (1991) found that “sensory declines” not only slowed down reading speed for text and led to extra sentence reading time, but also reduced participants’ free recall performance. They concluded that “sensory impairment can have significant secondary effects on higher level processes, such as memory, because it demands additional information processing capacity which becomes unavailable for inference, rehearsal and association” (p. 301). In the ageing literature, Schneider et al. (2000) reported that older adults could not recall as much detail as younger adults when required to answer comprehension questions after listening to discourse in conversational noise. Nevertheless, both age groups performed equally well after individual differences in hearing had been adjusted by increasing the signal-to-noise ratio (SNR) for the old, indicating that declines in hearing were a major contributor to elders’ poorer comprehension. Also, using eye tracking, Rayner et al. (2006) reported that younger adults’ eye movements while reading Old English font (difficult orthographic decoding) were very similar to those of the old while reading Times New Roman font (easy orthographic decoding). Finally, Pichora-Fuller et al. (1995, Experiment 2) showed that even though background noise differentially impaired elders’ performance, recall in both age groups was compromised in the severely noisy condition (at floor when SNR equalled zero).

Although these results are generally consistent with the effortfulness hypothesis, they often rely on measures of retrieval (i.e., recall) to make inferences about online decoding and semantic integration of auditory/visual stimuli. Lacking is direct evidence for whether there are age differences in noise effects on encoding processes during sentence comprehension.

Sentence processing requires surface-level decoding of the word form and lexical access (i.e., word-level processing) and deep-level conceptual integration (i.e., textbase-level processing; Kintsch, 1998). Word-level and textbase-level processes are conducted dynamically and interactively for mature and literate adults (Stine-Morrow et al., 2006). Word-level processes are to some extent prerequisite to textbase-level semantic analysis, but resources allocated to textbase-level processing are conducted more variably and are often correlated with immediate recall performance (Haberlandt, Graesser, Schneider, & Kiely, 1986; Stine-Morrow et al., 2006; Stine-Morrow, Noh, & Shake, 2010). To the extent that resources are not effectively allocated to textbase processing, the consequence would be a semantic representation of the text content that is lacking in fidelity.

A strong textbase representation may be manifested not only as high levels of recall in absolute terms, but also in the quality of recall. Textbase construction is a cyclical process that occurs iteratively as a reader moves through the text. An effective reader reactivates core ideas in an effort to draw connections across the text, so as to represent the meaning of the text as an organized and coherent whole. Because core ideas are reactivated, they are more memorable (Miller & Kintsch, 1980). In fact, the memorability of particular ideas is reliable (Rubin, 1985; Stine & Wingfield, 1988). Conditions that increase the difficulty of text processing (e.g., informational complexity, unfamiliarity) have been shown to differentially reduce the recall performance for the more memorable ideas among older adults (Hartley, 1993; Stine & Wingfield, 1988, 1990), suggesting an effect of these conditions on text organization.

Allocation to decode the surface form and allocation for semantic analysis are typically correlated in nature (Stine-Morrow et al., 2008), but these processes can be dissociated by orienting tasks that manipulate levels of processing (Craik & Lockhart, 1972), which can be thought of as the more general cognitive principle underpinning the effortfulness hypothesis (e.g., Rosenberg & Schiller, 1971). There is some evidence with younger samples that external noise can impair semantic processing in language understanding. For example, simulated visual impairment (e.g., changes in print size or contrast) slows down word-level processing (Bowers & Reid, 1997; Bullimore & Bailey, 1995), but importantly has secondary effects on semantic analysis in sentence understanding. By systematically altering the auditory signal of the sentence (i.e., low-pass filtering), Aydelott and Bates...
differences in the effects of visual noise on sentence processing have used offline measures of performance (word identification or sentence recall), providing no direct measures of the noise effects on comprehension processes. In the current study, we used the same paradigm as Gao et al. (2011) to examine whether internal and external noise have cumulative impact on reading among younger and older adults. Because the effortfulness hypothesis suggests that external and internal noise function similarly, it was hypothesized that ageing would exaggerate the adverse effect of environmental noise on word-level and textbase-level resource allocation as well as text recall. Thus, we predicted that (a) with increased levels of visual noise, effort allocated to word-level processes would increase at the expense of effort to semantic integration (textbase processes), especially for the old; and that (b) subsequently, the decreased allocation to semantic processing would compromise recall performance.

EXPERIMENT 1

Method

Participants

Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$, $p < 0.01$. Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$, $p < 0.01$. Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$, $p < 0.01$. Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$, $p < 0.01$. Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$, $p < 0.01$. Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$, $p < 0.01$. Older adults were recruited from the local community of Urbana-Champaign, and young adults were recruited from the Educational Psychology subject pool at the University of Illinois. Older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of subjects ($n = 31$ for the old; $n = 32$ for the young; mean age, old, $M_o = 69.5$ years, $SD = 7.2$; mean age, young, $M_y = 23.8$ years, $SD = 4.0$) did not differ in terms of years of education ($M_o = 16.0$, $SD = 2.6$; $M_y = 16.2$, $SD = 1.9$), $t < 1$, or vocabulary knowledge ($M_o = 51.3$, $SD = 8.5$; $M_y = 53.1$, $SD = 8.1$; Vocabulary subscale of the Wechsler Adult Intelligence Scale–Revised, WAIS–R, Wechsler, 1981), $t < 1$. Younger adults generally performed better on working memory span tasks (the average of loaded listening span and loaded reading span; Stine & Hindman, 1994; $M_o = 4.3$, $SD = 1.0$; $M_y = 5.7$, $SD = 1.2$), $t(61) = 4.99$,
Both groups of participants had normal or corrected-to-normal vision on both Snellen and Rosenbaum eye charts (as measured by ratio scores), but younger adults generally had better visual acuity than older adults on the Snellen, $M_r = 0.93, SD = 0.11; M_r = 0.99, SD = 0.05, t (61) = 2.53, p < .05$, and Rosenbaum, $M_r = 0.94, SD = 0.12; M_r = 1.00, SD = 0.00, t (61) = 3.15, p < .01$, test of visual acuity. An additional three older and three younger adults were tested but their data were removed because their reaction times (RTs) could not be modeled (i.e., regression of RTs onto text variables was not significant, $p < .01$, for two or more conditions). Furthermore, two older adults and one younger adult were classified as outliers on composite resource allocation indices (i.e., $z$ scores, as described later), based on a criterion of a $p < .01$ on the chi-square distribution of Mahalanobis distance (cf., Mertler & Vannatta, 2005).

**Materials and design**

Text materials for this study were three sets of twenty-four 18-word unrelated test sentences dealing with diverse topics in science, nature, and history, which were adapted from those used by Stine-Morrow et al. (2001). These sets of test sentences were equated in terms of word length, syntactic complexity, and propositional density (cf. Stine-Morrow et al., 2001, for details). Single sentences were chosen so that we could assess textbase processing in which discourse context was minimized, as our primary concern in this study was how visual noise affected the semantic analysis of sentence. Each test sentence was followed by another (short) sentence that provided a sensible continuation of the first (e.g., “In many species it is the females who shape evolution through their subtle exercise of choice in mating. They often choose mates who are bolder or brightly colored.”). The purpose of the continuation filler sentence was to guarantee that (1) the wrap-up effect at the end of the first sentence was not overestimated due to retrieval planning (as required by the recall task), and that (2) recall of content from the target sentence reflected semantic retrieval instead of merely surface-form rehearsal. Neither reading time nor recall was analyzed for these fillers.

Each word in the test sentence was coded in terms of word-level and textbase-level linguistic features that reflected different aspects of sentence processing. The word-level features encompassed the number of syllables and log word frequency (Francis & Kucera, 1982) of each word in the sentence, to estimate the time allocated to orthographic decoding and to word meaning access, respectively. Textbase-level features contained in the regression were whether the word was a newly introduced concept in the sentence (dummy-coded as 0/1) and cumulative conceptual load at sentence boundaries, calculated by multiplying the total number of concepts introduced in the sentence by the dummy-coded variable for the sentence-final word, to estimate the time allocated to immediate processing of new conceptual information and to the conceptual integration at sentence wrap-up, respectively (see Stine-Morrow et al., 2008, for details). These four text variables were chosen and were then clustered into word-level and textbase-level sentence processing, because of their strong and reliable effects on reading times.

Dynamic visual noise was generated, and text was projected simultaneously on a single monitor (iMac 17" LCD monitor 1,440 × 900 pixels with 32-bit colour, OS 10.4.10) using Matlab software. Level of visual noise was varied by assigning a randomly selected proportion of pixels (from both the text and the background) to a new randomly selected greyscale value after each refresh ($0.3 = $ low noise, $0.5 = $ medium noise, $0.7 = $ high noise), creating three levels of visual noise. Passages (embedded with varying levels of noise) were presented in the *moving window* paradigm (Just, Carpenter, & Wooley, 1982). The mean luminance of the iMac monitor was 82 cd/m²; the reading distance was maintained at 44 cm by a chin rest.

The three sets of 24 sentences were counterbalanced across three noise conditions (low, medium, high). Sentences within each noise condition were blocked for presentation, and within each block, sentences were randomized in a single fixed order for all participants. The order of noise
conditions was counterbalanced across subjects. Together this created nine unique combinations of sentence sets and noise levels. This design assured that each participant read all three sets of sentences and only read one set of materials at each of three levels of visual noise. Passages were presented using a nonproportional font (Courier, size 28).

A speeded lexical decision task was administered prior to the main experiment to insure that participants could identify isolated words in the noisy background. Three lists containing words with a wide range of word frequency (1–411 occurrences) and word length (3–9 letters) were generated and were counterbalanced across noise conditions. Participants only saw one list of words at each level of noise.

Procedure
Participants were tested individually in an experimental session that lasted 90 min on average. Self-reported overall levels of health, hearing, and vision were collected. Before the reading task, participants were administered the Vocabulary subscale of the WAIS–R; after completing the reading task, they were tested on the loaded Reading Span task, the loaded Listening Span task, and the vision acuity charts (Snellen and Rosenbaum).

Participants completed the lexical decision task, at the beginning of which they were told that some of the words would be “presented with some distractions like the static or snow on a fuzzy television picture”. After that, participants read the experimental texts in a self-paced moving window fashion and were encouraged to read as naturally as possible and “to remember as much of the information from the passages as possible”, because they would be occasionally asked to recall some of these passages. A “READY?” signal was presented in the centre of the screen at the beginning of each trial. The participants pressed the space bar, which removed this signal and triggered a fixation point (+) at the top left corner of the screen, indicating the spot where the first word of the passage would appear. The configuration of the entire two-sentence passage was indicated by dashes and punctuations that followed the fixation point. The next key press caused the first word to appear, and with another key press, the first word would be replaced with dashes, and the second word appeared where corresponding dashes had been. The remainder of the passage was triggered by successive key presses.

Participants were presented with four practice trials in each noise condition to make them familiar and comfortable with the procedure. Furthermore, an additional passage was included at the beginning of each of the three blocks as a warm-up trial and was not included in any analysis. After a randomly selected third of the trials (8 out of 24 in each block), the phrase “PLEASE RECALL NOW” appeared on the screen, signalling that the participant should recall aloud what was remembered from the passage. Recall protocols were recorded and later transcribed and scored using a gist criterion for propositional recall. Two independent raters scored a subset of protocols (from three randomly chosen participants of each age group) with good reliability \((r = .95)\), and then the remaining protocols were scored by either rater.

Results
Lexical decision task
Accuracy was above 95% in all conditions for both age groups, which assured us that noise did not impair readers’ ability to identify isolated words.

Reading time
Raw reading times were trimmed such that reading times within individual subjects in each noise condition that were greater than five standard deviations above the mean were replaced with that upper limit. This resulted in replacements of less than 0.74% of the raw data, and all the following analyses were based on the trimmed data. Older adults’ reading times were longer \((M_{old} = 703 \text{ ms}, SE = 44, M_{young} = 513 \text{ ms}, SE = 43, F_1(1, 61) = 9.54, p < .01; F_2(1, 430) = 48.37, p < .001)\); noise tended to increase reading time \((M_{low} = 594 \text{ ms}, SE = 32, M_{medium} = 602 \text{ ms}, SE = 34, M_{high} = 628 \text{ ms}, SE = 31), F_1(2, 122) = 1.99, \)
patterns of resource allocation

Reading times for each participant in each noise condition were decomposed into attentional resources (time) allocated to linguistic computations needed in sentence processing by regressing them onto the word-level and textbase-level features (Lorch & Myers, 1990). As described above, (a) the number of syllables and (b) log word frequency were used to isolate word-level processes; and (a) the dummy-code for new concepts and (b) the cumulative number of new concepts at sentence boundaries were used to isolate textbase processes. All the item-level predictors (i.e., text characteristics) were simultaneously entered to model the resource allocation to specific text processes for each individual within each noise condition. Such coefficients are reliable indicators of readers’ allocation policies across time (Stine-Morrow et al., 2001) and across texts (Stine-Morrow et al., 2008) and could be used to create composites that operationalize our theoretical constructs (Stine-Morrow, Gagne, Morrow, & DeWall, 2004).

The coefficients from these individual regressions were trimmed such that regression coefficients greater than 2.5 standard deviations of the group mean within each condition were replaced with that mean. This resulted in replacements of fewer than 3.25% of the regression coefficients for the old and 2.78% for the young. Means of individual parameters for each age group across noise conditions are shown in Table 1. Each of the word-level and textbase-level features was a reliable predictor of reading times at each level of noise, except for new concepts in the low- and medium-noise condition for the young.

We computed composite scores for word-level and text-level processes using the standardized z scores of each parameter across noise conditions within each age group, which are more reliable than individual parameters (e.g., Gao et al., 2011; Stine-Morrow et al., 2004). These composites were obtained by averaging z scores of the coefficients for syllable and word frequency, and for new concepts and sentence boundary, respectively. An alternative approach to analysis would be multi-level modelling in which reading times are simultaneously modelled for both subject and items effects (e.g., Baayen, Davidson, & Bates, 2008; Richter, 2006). However, because the effortfulness

\[ p = .14; \ F_2(2, 860) = 10.07, \ p < .001; \] but it did not interact with age, \( F_1 \) and \( F_2 < 1. \)

Table 1. Mean word RTs and allocation parameters as a function of visual noise for older and younger adults in Experiment 1

<table>
<thead>
<tr>
<th>Population</th>
<th>Measures of resource allocation</th>
<th>Low noise</th>
<th>Medium noise</th>
<th>High noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta )</td>
<td>SE</td>
<td>( \beta )</td>
<td>SE</td>
</tr>
<tr>
<td>Older adults</td>
<td>Word RT</td>
<td>689</td>
<td>45</td>
<td>686</td>
</tr>
<tr>
<td></td>
<td>Allocation parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Syllable</td>
<td>48***</td>
<td>10</td>
<td>40***</td>
</tr>
<tr>
<td></td>
<td>Word frequency</td>
<td>-18***</td>
<td>4</td>
<td>-23***</td>
</tr>
<tr>
<td></td>
<td>New concepts</td>
<td>89***</td>
<td>16</td>
<td>68***</td>
</tr>
<tr>
<td></td>
<td>Sentence boundary</td>
<td>111***</td>
<td>24</td>
<td>92***</td>
</tr>
<tr>
<td>Younger adults</td>
<td>Word RT</td>
<td>499</td>
<td>44</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td>Allocation parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Syllable</td>
<td>19**</td>
<td>9</td>
<td>23**</td>
</tr>
<tr>
<td></td>
<td>Word frequency</td>
<td>-17***</td>
<td>4</td>
<td>-19***</td>
</tr>
<tr>
<td></td>
<td>New concepts</td>
<td>16</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Sentence boundary</td>
<td>104***</td>
<td>24</td>
<td>76***</td>
</tr>
</tbody>
</table>

Note: RTs (reaction times) and allocation parameters in ms. Standard errors in parentheses. Parameters were compared to zero in single-sample \( t \) test.

\( ^* p < .05. ^{**} p < .01. ^{***} p < .001. \)
hypothesis predicts that increased noise will increase word-level processing and decrease textbase-level processing, and that ageing will exaggerate these effects, this analytic approach would have constrained our test of hypothesis to individual indicators and would not have allowed us to test our hypotheses at the construct level. Therefore, we computed and tests effects in our theory-driven composites.

The three-way interaction of age (young and old), noise (low, medium, high), and level of processing (word and textbase) was significant, $F(2, 122) = 3.46$, $p < .05$. This interaction plotted in Figure 1 showed that for the older adults, but not for the young, noise produced the predicted dissociation of word and textbase processing, $F(2, 60) = 6.19$, $p < .01$, for the old; $F < 1$ for the young. For the older group only, the noise evoked attention to word-level processing, $F(2, 60) = 3.92$, $p < .05$, simultaneously depleting attentional resources available for textbase-level processing, $F(2, 60) = 3.32$, $p < .05$.

Recall performance
Test sentences were decomposed into component propositions using the Kintsch system (e.g., Miller & Kintsch, 1980). A propositional analysis (Turner & Greene, 1978) of the 24 sentences yielded 179 propositions, which contribute to the meaning of the sentence. Two independent raters scored three protocols that were randomly selected from each age group, with excellent consistency ($r = .95$, $p < .01$); the remaining recall protocols were scored by either rater. We analysed the recall in a crossed-random-effects hierarchical linear model (Baayen et al., 2008), in which age (old, young), noise (low, medium, high), and their interaction were treated as fixed effects while the between-subject and between-item variations were simultaneously treated as crossed-random effects. Noise did not reduce overall recall ($M_{\text{low}} = 57.1\%$, $SE = 2\%$; $M_{\text{medium}} = 56.2\%$, $SE = 2\%$; $M_{\text{high}} = 57.3\%$, $SE = 2\%$), $z = 1.11$, $p = .27$, nor did age ($M_{o} = 55.8\%$, $SE = 3\%$; $M_{y} = 58.0\%$, $SE = 3\%$), $z < 1$. The age by noise interaction was also not significant, $z = 1.46$, $p = .15$. For both age groups, overall recall remained unaffected by the external noise.

To test whether noise impacted the quality of recall, we analysed recall for only the top 20% most memorable propositions (based on the probability of recall for each proposition in the low-
noise condition, collapsing across subjects; Rubin, 1985; Stine & Wingfield, 1988) in a separate crossed-random effects hierarchical linear model. Consistent with the integration-disruption hypothesis, noise impaired the recall of the core propositions, \( z = 3.57, p < .001 \). As shown in Figure 2, this effect was moderated by age in a significant noise by age interaction, \( z = 3.17, p < .01 \), such that noise differentially reduced the recall of important ideas in sentences for the old, \( z = 3.54, p < .001 \), without compromising recall for the young, \( z < 1 \). Thus, the pattern of recall was in good alignment with the effects of visual noise on resource allocation during reading.1

However, there were individual differences among the old in the extent to which visual noise disrupted semantic processing and recall. We capitalized on this variability to examine the direct relationship between noise-induced textbase neglect and reduced recall in the high-noise condition (relative to the low-noise condition), reasoning that participants whose recall was more decreased by noise would also have shown exaggerated effects of noise on resource allocation compared to those who were not. This was, indeed, the case. Participants whose recall in high noise fell below one standard error of measurement (\( SEM \)) of that in low noise were characterized as the decline group (\( N = 20 \)), whereas those whose recall fell above this statistical criterion were characterized as the stable group (\( N = 11 \); cf. Schaie, 1984). Relative to the low-noise condition, in the high-noise condition the decline group allocated more resources to word-level processes (\( M_{\text{low}} = .04, SE = .14, M_{\text{high}} = .43, SE = .18 \), \( t(19) = 2.13, p < .05 \), and less time to textbase construction (\( M_{\text{low}} = .31, SE = .23, M_{\text{high}} = .09, SE = .17 \), \( t(19) = 2.04, p < .05 \)). However, the stable group did not show change in resource allocation across noise conditions for word processes (\( M_{\text{low}} = -.37, SE = .19, M_{\text{high}} = -.20, SE = .24 \), \( t < 1 \), or to textbase processes (\( M_{\text{low}} = -.07, SE = .31, M_{\text{high}} = -.24, SE = .23 \), \( t(10) = 1.09, p = .30 \).

**Discussion**

For the older adults only, there was an interaction between visual noise and level of processing on resource allocation in reading. In the face of a degenerated linguistic signal, older readers adjusted their resource allocation so as to encode the surface form of the text, but at the cost of more superficial textbase processing. This pattern is consistent with the idea that age-related deficits at both the sensory level and the cognitive level (“internal noise”) exacerbate the effects of external noise to decrease the resources available for semantic analysis. As a consequence of this more superficial semantic analysis, the quality of their recall was reduced, resulting in the retention of relatively fewer central ideas. These results represent a substantive replication of previous findings of noise effect on offline memory performance (e.g., Dickinson & Rabbitt, 1991; Gao et al., 2011) and support the effortfulness hypothesis.

**EXPERIMENT 2**

Unlike in Gao et al. (2011), allocation policies and recall quality for the young were relatively unaffected by noise in the context of the first

---

1 Results were substantively the same when we performed this analysis with the top 15% or top 10% most memorable propositions, with significant age by noise interactions in both cases, \( z = 3.57, p < .001 \), for the top 15%; \( z = 3.94, p < .001 \), for the top 10%.
experiment. It is plausible that the noise level had not been intense enough to tax their sentence understanding, leaving their performance at ceiling. However, the effortfulness hypothesis would also predict that by increasing external noise, the younger adults would function similarly to the older adults. Thus, this experiment was a replication of the first but with a more extreme manipulation of visual noise.

Method

Participants
Eighteen older adults ($M_o = 67.9$, $SD = 4.9$) and 18 young adults ($M_y = 21.5$, $SD = 4.4$) were recruited as in the first experiment. As before, older participants were screened for neurological disorders and learning disabilities in a telephone interview prior to scheduled for testing. The two groups of participants did not differ in terms of years of education ($M_o = 15.2$, $SD = 2.7$; $M_y = 14.7$, $SD = 1.7$), $t(34) = .64$, $p = .53$, or vocabulary knowledge ($M_o = 46.6$, $SD = 6.9$; $M_y = 45.3$, $SD = 5.3$) $t(34) = 0.62$, $p = .54$. Younger adults outperformed older adults on the working memory span task ($M_o = 4.3$, $SD = 0.9$; $M_y = 5.5$, $SD = 1.2$), $t(34) = 3.55$, $p = .001$. Younger and older adults’ corrected visual acuity were comparable, $M_o = .97$, $SD = .09$; $M_y = .98$, $SD = .06$, $t < 1$ for Snellen eye chart, and $M_o = .95$, $SD = .10$; $M_y = .99$, $SD = .04$, $t(34) = 1.52$, $p = .14$ for Rosenbaum eye chart. Participants were statistically screened as in the first experiment, but no data were removed.

Materials, design, and procedure
The parameters for visual noise densities were set to .3, .6, and .8 for low-, medium- and high-noise conditions, respectively. All the other aspects were the same as those in Experiment 1.

Results

Lexical decision task
Accuracy was above 95% in all conditions for both age groups, suggesting that the ability to identify isolated words was not compromised in noise.

Reading time
Raw reading times were trimmed based on the same criteria as those employed for Experiment 1, which led to the replacement of less than 0.64% of the raw data. The main effects of age, $F_1(1, 34) = 8.53$, $p < .01$, $F_2(1, 430) = 57.72$, $p < .001$; $M_{old} = 750$ ms, $SE = 60$, and $M_{young} = 506$ ms, $SE = 60$) and noise, $F_1(2, 68) = 25.39$, $p < .001$, $F_2(2, 860) = 78.64$, $p < .001$; $M_{low} = 583$ ms, $SE = 61$, $M_{medium} = 612$ ms, $SE = 59$, and $M_{high} = 690$ ms, $SE = 61$) on reading time were significant. As in Experiment 1, visual noise did not interact with age in influencing overall word-by-word reading time, $F_1$ and $F_2 < 1$ (see Table 2). It was noteworthy that the effect of noise on reading time was more reliable in this experiment than in the first, suggesting that the visual noise was more taxing.

Patterns of resource allocation
Using the same trimming criterion as that described in Experiment 1, 1.39% of regression parameters were replaced for the old group, and 2.78% were replaced for the young group. Means of trimmed individual parameters for each age group across noise conditions are shown in Table 2. All the selected word- and textbase-level features were reliable predictors of reading times across noise conditions (except for new concepts in medium and high noise).

In the three-way (age, noise, and levels of sentence processing) analysis of variance (ANOVA) on composite scores for word-level and text-level processing, a strong interaction of noise and level of sentence processing, $F(2, 68) = 16.82$, $p < .001$ (see Figure 3) that did not vary with age, $F < 1$, was found. The noise effects were marginally significant for word-level processing for both older, $F(2, 34) = 2.66$, $p = .08$, and younger, $F(2, 34) = 3.17$, $p = .06$, adults, but were highly reliable for textbase-level processing, $F(2, 34) = 5.36$, $p < .01$, for the old and $F(2, 34) = 5.40$, $p < .01$, for the young. For both groups of participants, noise affected word-level processing by requiring extra attention to orthographic decoding and lexical access and therefore by reducing the attentional resources that could have been used for textbase-level processing.
Recall performance

As in Experiment 1, in the hierarchical crossed-random-effects model, noise, age, or their interaction did not impact the quantity of recall. However, consistent with the patterns of resource allocation, noise impaired the recall of the core (the 20% most memorable) propositions, \( z = 2.24, p < .05 \), and, as shown in Figure 4, this effect did not interact with age, \( z < 1 \). Visual noise reduced the recall of important ideas in

\[ \text{Table 2. Mean word RTs and allocation parameters as a function of visual noise for older and younger adults in Experiment 2} \]

<table>
<thead>
<tr>
<th>Population</th>
<th>Measures of resource allocation</th>
<th>Low noise</th>
<th>Medium noise</th>
<th>High noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta )</td>
<td>SE</td>
<td>( \beta )</td>
<td>SE</td>
</tr>
<tr>
<td>Older adults</td>
<td>Word RT</td>
<td>697</td>
<td>311</td>
<td>728</td>
</tr>
<tr>
<td>Allocation</td>
<td>Syllable</td>
<td>42**</td>
<td>11</td>
<td>44***</td>
</tr>
<tr>
<td></td>
<td>Word frequency</td>
<td>-30***</td>
<td>5</td>
<td>-35***</td>
</tr>
<tr>
<td></td>
<td>New concepts</td>
<td>45**</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Sentence boundary</td>
<td>58***</td>
<td>9</td>
<td>44***</td>
</tr>
<tr>
<td>Younger adults</td>
<td>Word RT</td>
<td>468</td>
<td>188</td>
<td>495</td>
</tr>
<tr>
<td>Allocation</td>
<td>Syllable</td>
<td>23**</td>
<td>11</td>
<td>26**</td>
</tr>
<tr>
<td></td>
<td>Word frequency</td>
<td>-14**</td>
<td>5</td>
<td>-19***</td>
</tr>
<tr>
<td></td>
<td>New concepts</td>
<td>26*</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Sentence boundary</td>
<td>40***</td>
<td>9</td>
<td>33**</td>
</tr>
</tbody>
</table>

\( \text{Note: RTs (reaction times) and allocation parameters in ms. Standard errors in parentheses. Parameters were compared to zero in single-sample t test.} \)

\* \( p < .05 \).
\** \( p < .01 \).
\*** \( p < .001 \).

\[ \text{Figure 3. Resource allocated to word-level and textbase-level processing as a function of visual noise for older and younger adults in Experiment 2.} \]

\[ \text{Recall performance} \]

\[ \text{As in Experiment 1, in the hierarchical crossed-random-effects model, noise, age, or their interaction did not impact the quantity of recall. However, consistent with the patterns of resource allocation, noise impaired the recall of the core (the 20% most memorable) propositions, } z = 2.24, p < .05, \text{ and, as shown in Figure 4, this effect did not interact with age, } z < 1. \]

\[ ^2 \text{Results were substantively the same when we performed this analysis with the top 15% or top 10% most memorable propositions with a significant effect of noise on recall, } z = 1.89, p = .05 \text{ for the top 15%, } z = 2.44, p < .05 \text{ for the top 10%, and no moderation of this effect by age, } z < 1 \text{ for both cases.} \]
Discussion

The intensity of the visual noise affected multiple levels of sentence processing. Both age groups allocated greater resources to decoding and lexical access, while decreasing resources allocated to conceptual integration to the same degree. Comparing these results to those of the first experiment, younger adults tolerated external noise to a more extreme level before adjusting their allocation of resources (cf. Gao et al., 2011). Within this range of external noise, younger and older readers were similarly affected in reducing attention to textbase construction, and both age groups produced text recall that was less likely to contain core ideas. The experimental manipulation of noise that resulted in both depressed attentional allocation during reading and less organized recall provides further direct evidence for the effortfulness hypothesis—for the first time empirically localizing the effect to encoding processes. Extreme external noise taxed word-level processes at the expense of attentional resources typically available for conceptual integration and as a consequence impaired the quality of recall.

GENERAL DISCUSSION

The data from the two experiments collectively are in alignment with the effortfulness hypothesis and show that external noise and internal noise can have cumulative effects in distracting from effective semantic analysis. When decoding of the surface form was challenged by a noisy presentation, word and textbase processes traded off for limited attentional resources. In the first experiment, in which noise was manipulated within a moderate range, there was selective impact on the resource allocation of older readers. This is consistent with the idea that internal noise exacerbated the effects of the external noise. Within this range, younger adults were relatively resilient. When noise intensity was relatively strong in the second experiment, resource allocation to semantic processing was reduced among both age groups because attention was diverted to decoding the surface form.

The quality of recall was impaired by noise for the old in the first experiment and for both groups in the second experiment. Our findings align with earlier studies (Hartley, 1993; Stine & Wingfield, 1988) in showing that difficult encoding conditions can exact a price on the quality of recall, even if not apparent in the overall amount of information recalled. The novel finding is that these effects can occur as a function of challenges in decoding the surface form that are completely unrelated to demands for semantic processing per se.

3 With the more extreme manipulation of noise in this experiment, 14 older and 15 younger participants showed reduced recall from low to high noise, and 4 older and 3 younger participants did not, using the criterion in Experiment 1. These smaller sample sizes did not allow a meaningful test of change in resource allocation for these two recall groups.

4 It would have been interesting to follow up these experiments with an even more extreme manipulation of visual noise. Unfortunately, our attempt to do this experiment failed. Piloting with a high-noise condition of .85 revealed that a fluent reading experience was impossible, with self-pacing becoming erratic as reading turned into a problem-solving activity (for some very unhappy research participants).
The findings that even mild levels of noise undermined the reading experience and outcome for the old highlighted the importance of considering sensory declines in cognitive ageing research. It is noteworthy that even though all the older adults in our study had vision that was in the normal range or corrected-to-normal range, numerically they had worse visual acuity than the young. Future studies should integrate sensory as well as fluid ability declines, situate them in the same framework, and dissociate their respective roles in a wide range of perceptual and cognitive tasks (Wingfield et al., 2005). Programmes designed for cognitive optimization for the older adults should render ideal levels of signal-to-noise ratio (SNR) when materials and/or testing environment are designed.

It is well documented that older adults are more resilient in creating a situation model representation of the text (i.e., spatial, temporal, emotional, causal relationships of the subjects, objects, or events) than are the young, which is in part due to elders’ increased verbal ability and world knowledge and the reduced demands that situation model features pose for encoding, storage, and retrieval in working memory (see Radvansky & Dijkstra, 2007, for a review). From a psycholinguistic perspective, it would be intriguing to examine the compensatory role of situation model on lower levels of textual processing as well as the adverse effect of perceptual noise on more resilient situation-level representation.

Our findings are also relevant to a recent debate in the literature about the possibility that learning can actually be enhanced by increasing the difficulty of encoding. Diemand-Yauman, Oppenheimer, and Vaughan (2011) reported that high-school and college students improved learning when reading was made disfluent with hard-to-read fonts (e.g., Bodoni MT, Comic Sans MS). They explained their findings in terms of Bjork’s notion of desirable difficulty (e.g., Schmidt & Bjork, 1992): The disfluency in decoding induced the allocation of resources and deeper processing to the learning task. We do not have a ready explanation for this discrepancy, but consider two possibilities.

First, unusual fonts may engage attention in part by introducing novelty and/or implying significance (e.g., as in italics to imply stress) and, as such, encourage deeper processing. Older readers have been found to slow down to an extent similar to younger adults when encountering an unusual font (Rayner et al., 2006), but we know of no research investigating this as a source of compensation for ageing readers. Our manipulation of disrupting reading fluency with manipulation of visual noise measurably disrupted processing of the surface form without any pragmatic signalling of novelty or importance. While we agree that engaging readers’ attention for processes related to textbase construction—a desirable difficulty—can improve performance (e.g., Stine-Morrow et al., 2010), we believe that the principles that fully define the nature of difficulty that are beneficial to learning remain to be specified. If nothing else, our results may sound a cautionary note to those who would take the Diemand-Yauman et al. (2011) findings as motivation to adopt muddy instructional materials for adult learners under the assumption that this will improve learning, a note immediately sounded in the popular press (Learning difficulties, 2010; Brooks, 2010).

Second, it may actually be the case that allocation of effort to the surface form and allocation to the textbase are yoked under some conditions. Interestingly, even though word and textbase processing can clearly be dissociated (as shown here, and also by Smiler, Gagne, & Stine-Morrow, 2003), in the normal ecology of reading, word-level and textbase allocation are moderately correlated (Stine-Morrow et al., 2008; Stine-Morrow et al., 2006), suggesting that readers may often adopt a heuristic that does not privilege any one level of analysis (i.e., enhanced reading engagement entails attentional allocation collectively to reading computations; cf. Stine-Morrow et al., 2008, for a discussion). This fact makes it entirely plausible that certain forms of challenge in processing the surface form may result in improved effort across the levels of language analysis. It is important for future research to determine what sorts of difficulties are truly desired and to define that “sweet spot.”
that engages attention productively without distraction.

REFERENCES


