

Processing anomalous anaphors

Anne E. Cook

Published online: 6 May 2014
© Psychonomic Society, Inc. 2014

Abstract Previous researchers have demonstrated that readers may engage in shallow, or incomplete, processing when the semantic overlap between current information and previously encountered information is high. The present study investigated whether these effects would occur during processing of unambiguous noun phrase anaphors, for which there was only a single possible antecedent. Participants read passages containing anaphors that were correct, incorrect but highly related, or incorrect and low-related, with respect to previously encountered information. The time required to process the anaphor was a function of the goodness of fit between the anaphor and the antecedent; anaphors that were incorrect but highly related to the antecedent were processed more quickly than those that were incorrect and low-related. This occurred regardless of the distance between the anaphor and the antecedent. However, reading times results from a spillover sentence indicated that readers subsequently validated the anaphor against the information in memory, resulting in continued processing difficulty for both the incorrect-high- and –low-related anaphor conditions. The results are consistent with a three-stage comprehension model in which information is activated, integrated on the basis of its goodness of fit with the contents of working memory, and then validated against information in long-term memory.

Keywords Discourse processing · Reading · Shallow processing

Early models of comprehension operated under the assumption that readers would not proceed to the next word or phrase

A. E. Cook (✉)
Educational Psychology Department, University of Utah,
1721 Campus Center Drive, SAEC 3220, Salt Lake City,
UT 84112, USA
e-mail: Anne.Cook@utah.edu

in a text unless comprehension of the preceding information was complete (e.g., Just & Carpenter, 1980). Other researchers, however, have provided evidence that readers may continue on in a text before they are finished processing the preceding information (e.g., Duffy & Rayner, 1990; Ehrlich & Rayner, 1983) and that they may not always “finish” processing at all. Terms such as readers’ standards of coherence (van den Broek, Risden, & Husebye-Hartmann, 1995), good-enough representations (Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007), and shallow processing (Sanford, 2002; Sanford & Emmott, 2012; Sanford & Graesser, 2006; Sanford & Sturt, 2002) have been used to describe situations in which readers appear to be satisfied with incomplete or even incorrect interpretations of text. However, there are a variety of ways in which incomplete processing may occur. For example, readers may not completely integrate activated content with incoming information at all, or this process may not be completed by the point at which processing is measured during a task. “Shallow processing”, therefore, is an umbrella term that can describe a variety of incomplete processing situations. The purpose of this study was to investigate one specific type of shallow processing—those situations in which readers appear to process incoming information on the basis of its goodness of fit (i.e., conceptual overlap) with the contents of working memory, resulting in incorrect or incomplete interpretations.

The factors that influence the degree to which readers’ processing will be “incomplete” or “shallow” were first investigated using question-answering tasks, such as the Moses illusion (e.g., Bredart & Docquier, 1989; Bredart & Modolo, 1988; Erickson & Mattson, 1981; Kamas, Reder, & Ayers, 1996; Reder & Cleeremans, 1990; Reder & Kusbit, 1991), in which readers responded to questions that contained semantic anomalies (e.g., “How many animals of each kind did Moses take on the Ark?”). In these tasks, readers failed to detect the anomaly as much as 40 % of the time when the anomaly (e.g.,

Moses) was highly related to the correct term (e.g., *Noah*). Similarly, Sanford, Sturt, and colleagues (e.g., Sanford, Sanford, Filik, & Molle, 2005; Sanford, Sanford, Molle, & Emmott, 2006; Sturt, Sanford, Stewart, & Dawydiak, 2004; Ward & Sturt, 2007) implemented a text change detection task, in which participants were asked to detect minor wording changes (e.g., changing *hat* to *cap*) between two presentations of the same text. Readers were less successful at detecting changes that represented small meaning changes (e.g., *hat* to *cap*), rather than large meaning changes (e.g., *hat* to *dog*). Researchers have also used an “incidental anomaly detection task,” in which readers are presented with short texts and are told that some of them may contain anomalies (Barton & Sanford, 1993; Bohan & Sanford, 2008; Daneman, Hannon, & Burton, 2006; Hannon & Daneman, 2004; Sanford, Leuthold, Bohan, & Sanford, 2011). In these tasks, readers frequently fail to detect anomalous noun phrases, such as *surviving dead* (e.g., Barton & Sanford, 1993) or *tranquilizing stimulants* (e.g., Hannon & Daneman, 2004), presumably because the contents of the noun phrases are highly related to preceding information in the passage. The consistent finding across all these types of tasks has been that readers often do not detect anomalies unless those anomalies reflect large changes in meaning and/or unless they have received additional attentional focus via syntactic or prosodic manipulations (Bredart & Docquier, 1989; Bredart & Modolo, 1988; Sanford et al., 2006; Sturt et al., 2004). Sanford and colleagues (Sanford, 2002; Sanford & Emmott, 2012; Sanford & Graesser, 2006; Sanford & Sturt, 2002) have interpreted such findings as evidence that readers may often engage in shallow processing, in which comprehension appears to proceed smoothly as long as the semantic match (or goodness of fit) between a text and previously encountered information is sufficient to meet some internal criterion. In addition, when this type of shallow processing does occur, it is typically assumed that readers move on in the text with no subsequent consequences for comprehension.

The present study investigates how readers will process anomalies in discourse when no explicit anomaly detection task is required. Several sentence-processing studies have demonstrated that readers will arrive at underspecified or incomplete representations when presented with complex syntactic structures (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Swets, Desmet, Clifton, & Ferreira, 2008; Traxler, Pickering, & Clifton, 1998). In addition, researchers have found that readers may arrive at underspecified or incomplete representations when faced with anaphors involving ambiguous pronouns (Greene, McKoon, & Ratcliff, 1992; Poesio, Sturt, Artstein, & Filik, 2006; Stewart, Holler, & Kidd, 2007; but see also Love & McKoon, 2011).

With direct noun phrase anaphora that require the reinstatement of an antecedent, however, some researchers have argued that it is necessary for readers to fully resolve anaphoric

references in order for comprehension to proceed (e.g., Graesser, Singer, & Trabasso, 1994; Singer, Graesser, & Trabasso, 1994), whereas other findings suggest that full resolution may not be necessary (e.g., Klin, Guzmán, Weingartner, & Ralano, 2006; Klin, Weingartner, Guzmán, & Levine, 2004; Levine, Guzmán, & Klin, 2000; see also Cook, Myers, & O'Brien, 2005; Cook & O'Brien, *in press*; Nieuwland & Van Berkum, 2008). For example, Levine et al. (see also Klin et al., 2006; Klin et al., 2004) presented participants with passages that contained an anaphoric noun phrase (e.g., *dessert*) that referred to a previously encountered antecedent (e.g., *tart*). When a same-category distractor concept (e.g., *cake*) was elaborated in the text after the presentation of the antecedent, but before the presentation of the anaphor, readers failed to reactivate and reinstate the appropriate antecedent (e.g., *tart*) when they encountered the anaphor (e.g., *dessert*). As compared with a control condition, there was no response time advantage for the antecedent (e.g., *tart*) immediately after reading the anaphoric phrase, nor was there a reading time advantage for a subsequent sentence that explicitly reinstated *tart* as the antecedent. Klin et al. (2006; Klin et al., 2004; Levine et al., 2000) argued that the signal from the anaphor was divided between the correct antecedent (e.g., *tart*) and the related distractor (e.g., *cake*), so activation for the antecedent itself did not reach a critical threshold, and thus the antecedent was never fully activated or reinstated. More important, Klin et al. (2006) suggested that although the specific lexical item representing the antecedent was never activated, readers still partially encoded the anaphoric inference. Klin et al. (2006) offered two explanations for this partial encoding hypothesis: Readers may have activated only a subset of conceptual information about the antecedent without reactivating the specific lexical item, or readers may have reactivated enough information from the earlier context that integration of the anaphoric phrase could proceed without actual activation of the specific lexical item itself (see Cook, Limber, & O'Brien, 2001, for direct evidence of inference activation without accessing specific lexical information).

In another study of anaphoric processing, O'Brien and Albrecht (1991) demonstrated that readers may not arrive at the appropriate anaphoric inference if there is sufficient information in the text to support an alternative, but incorrect, conclusion. Their participants read passages that contained an anaphoric reference to a concept explicitly stated in the text (e.g., *cat*). However, if the text also contained information relating to an unmentioned concept (e.g., *skunk*), that unmentioned concept was reactivated in place of the appropriate, explicitly mentioned antecedent. Thus, the amount of contextual support for an antecedent can also influence whether or not readers engage in complete or shallow processing of an anaphor.

In the anaphoric processing studies just described, the anaphor was ambiguous, there was a concept that served as

a distractor to the antecedent, or the context supported an incorrect antecedent. According to Klin et al. (2006; Klin et al., 2004; Levine et al., 2000), the antecedent was not fully reactivated under these conditions, so resulting resolution of the anaphor was incomplete. The first question addressed here is how readers will process anomalous anaphors that are unambiguous and for which no distractor antecedents are present. Consider the first sample passage in the Appendix. The protagonist, Terry, goes to a music shop and buys an instrument, which is described over several sentences. This information is backgrounded and then later reinstated when Terry shows a friend the *cello* she bought. In the correct condition, the anaphor, *cello*, is a correct reference to the previously mentioned antecedent, *cello*. On the basis of previous research (O'Brien, Raney, Albrecht & Rayner, 1997), readers should quickly activate and reinstate the antecedent, leading to short reading times on the sentence containing the anaphor. The correct condition served as a control for two anomalous anaphor conditions. In the two anomalous conditions, the anaphor, *cello*, is incorrect but highly related to the actual antecedent (*violin*, in the incorrect–high-overlap condition), or it is incorrect and not highly related to the actual antecedent (*oboe*, in the incorrect–low-overlap condition). Given that there is only one possible antecedent present in the text and the anaphor is not ambiguous, the antecedent should be relatively easy to reactivate, because there is no competition from a potential alternative antecedent. If so, resolution of the anaphor is not likely to be shallow or incomplete under these conditions; readers should experience difficulty processing the anomalous anaphor in the reinstatement sentence in both incorrect–overlap conditions. That is, according to a *full-resolution account*, reading times on the reinstatement sentence should be shorter in the correct condition than in either the incorrect–high- or incorrect–low-overlap conditions, and there should be no difference between the two incorrect–overlap conditions. However, there is strong empirical support from the anomaly detection literature and from the previously mentioned studies on anaphora for the hypothesis that readers may not fully resolve an anaphor even under conditions in which relevant information (e.g., the antecedent) is highly accessible, as long as the semantic overlap between the anaphor and that previous information is strong. According to this view, initial integration of the anaphor is based on its goodness of fit with the information in working memory. Working memory would presumably include activated contextual and/or conceptual information about the antecedent; *contextual information* refers to information about the antecedent previously encountered in the passage, whereas *conceptual information* refers to information related to the antecedent activated from general world knowledge. If initial integration of an anaphor is based on goodness of fit with the contents of working memory, reading times on the reinstatement sentence should correspond to the degree of conceptual

overlap between the anaphor and the antecedent. Thus, the *goodness-of-fit account* would predict shorter reading times in the correct condition than in the incorrect–high-overlap condition and shorter reading times in the incorrect–high-overlap condition than in the incorrect–low-overlap condition.

The second issue investigated in this study concerned processing after readers have moved on from the reinstatement sentence. Given Duffy and Rayner's (1990; see also Ehrlich & Rayner, 1983) finding that readers may continue to process an anaphor after they move on in the text, it is also important to examine processing on a subsequent "spillover" region of text; this was done in all experiments reported here. If readers fully process the anaphor upon encountering it, there should be no subsequent effects of antecedent condition on this spillover region. In this case, there should be no lingering difficulty due to anomalous anaphors; reading times on the sentence immediately following the anaphor should not differ for the correct, incorrect–high-overlap, and incorrect–low-overlap conditions. Whether the goodness-of-fit account would make a different set of a priori predictions regarding subsequent processing effects is not clear from the previous literature.

A final issue investigated here was whether factors known to influence antecedent reinstatement will also influence how anomalous anaphors are processed. For example, antecedents that are unelaborated or that are more distant from the anaphor are not reactivated or reinstated as quickly as those that are elaborated or more recent (O'Brien, 1987; O'Brien, Albrecht, Hakala, & Rizzella, 1995; O'Brien, Plewes & Albrecht, 1990; O'Brien et al., 1997). Although these factors have been examined with respect to antecedent retrieval, they have not been investigated in the context of anomalous information. In the present study, anomalous anaphors may be processed differently when the amount of intervening text between the antecedent and the anaphor is reduced and/or the antecedent has been elaborated in the text. These issues were investigated in Experiments 2 and 3.

Experiment 1

Method

Participants and design

Thirty University of New Hampshire undergraduates enrolled in introductory psychology courses participated in exchange for partial course credit. All participants were native speakers of English.

Antecedent condition was manipulated within subjects, and there were three levels: correct, incorrect–high-overlap, and incorrect–low-overlap. The data for the reinstatement and spillover sentences were analyzed separately.

Materials

To ensure that the two incorrect antecedents did, in fact, differ in overlap with the anaphor, a rating experiment was conducted. Twenty-six University of New Hampshire students who did not participate in any other experiments reported here were asked to rate the similarity of a pair of concepts on a 5-point scale (where 1 = *extremely dissimilar* and 5 = *extremely similar*). Each word pair consisted of the anaphor and either the incorrect–high-overlap antecedent or the incorrect–low-overlap antecedent. Participants rated the anaphor/incorrect–high-overlap antecedent pairs as significantly more similar (mean = 3.75, $SD = 0.36$) than the anaphor/incorrect–low-overlap antecedent pairs (mean = 2.61, $SD = 0.85$), $F_1(1, 24) = 71.86$, $MSE = .24$, partial $\eta^2 = .75$; $F_2(1, 22) = 170.55$, $MSE = .09$,¹ partial $\eta^2 = .89$.

The materials for Experiment 1 were 24 passages similar to the first example in the Appendix. Each passage began with an introductory section (mean = 25.58 words) that served to introduce the protagonist of the story. This was followed by an elaboration section that introduced the antecedent and described its distinctive features (e.g., cellos are large and heavy, you sit down to play them, etc.). The antecedent was always mentioned once explicitly and twice implicitly. The antecedent was a correct match to a later-encountered anaphor, incorrect but highly related to the anaphor, or incorrect and low-related to the anaphor. The mean lengths of the elaboration sections for the correct, incorrect–high, and incorrect–low conditions were 84.83, 84.54, and 84.71 words, respectively. The elaboration section was followed by a background section (mean = 75.63 words); this section was included to ensure that the antecedent was no longer active in memory when the anaphor was presented. The background was followed by a reinstatement sentence that included a direct anaphoric reference (e.g., *cello*) to the antecedent. Note that this sentence was the same across all three conditions. A spillover sentence followed the reinstatement sentence. The mean lengths of the reinstatement and spillover sentences were 38.21 and 37.63 characters, respectively. A closing section (mean = 16.13 words) ended the passage. Each passage was followed by a yes/no comprehension question designed to ensure that participants were reading carefully; questions did not refer to the antecedent or anaphor. There were an equal number of “yes” and “no” questions.

Eight filler passages were included to ensure that there were equal numbers of passages with and without anomalous anaphors. Filler passages were similar in length and structure to the experimental passages. In addition, for the filler

passages, there were an equal number of “yes” and “no” comprehension questions.

Three materials sets were constructed and counterbalanced such that an equal number of passages appeared in each condition, and across the materials sets, each passage appeared once in each condition.

Procedure

Participants were randomly assigned to one of the materials sets. Each participant was run individually in a session that lasted approximately 30 min. All materials were presented on a monitor controlled by a computer in an adjacent room.

Participants were instructed to rest their right thumbs on a line-advance key, their right index fingers on a “yes” key, and their left index fingers on a “no” key. Each trial began with the word “READY” in the middle of the screen. When participants were ready to read a passage, they pressed the line-advance key. Each press of the key erased the current line and presented the next line. Reading time was measured as the time between keypresses. Each participant was instructed to read at a comfortable, normal reading pace. After the last line of the passage disappeared from the screen, the cue “QUESTION” appeared in the middle of the screen for 2,000 ms, followed by the comprehension question. Participants were instructed to respond to the comprehension question by pressing either the “yes” or the “no” key. Participants were also instructed that answering the comprehension questions was the most important part of the experiment and that they should respond as quickly as possible without sacrificing accuracy. On the trials on which participants made errors, the word “ERROR” appeared in the middle of the screen for 750 ms. Before reading the experimental passages, participants read three practice passages to ensure that they were thoroughly familiarized with and understood the procedure.²

Results and discussion

Comprehension question accuracy rates were quite high (over 80 % for all participants), and there were no differences in accuracy as a function of condition in Experiment 1, $F_s < 1$.

¹ In all analyses reported, F_1 always refers to tests against error terms based on participants’ variability, and F_2 always refers to tests against an error term based on items’ variability. All analyses were significant at the standard alpha level of .05, unless otherwise indicated.

² To ensure that readers did not become more sensitive to and, thus, develop strategies for processing anomalous anaphors over the course of the experiment, reading times for passages that appeared in the first quartile of both experiments were compared with reading times for passages that appeared in the last quartile of the passages. The sizes of the differences between conditions did not differ significantly from the first to the fourth quartile in Experiment 1, $p_s > .12$.

Table 1 Mean reading times (and standard deviations) in milliseconds for the reinstatement and spillover sentences as a function of condition in Experiment 1

	Correct	Incorrect–High-Overlap (IH)	Incorrect–Low-Overlap (IL)	<i>IH vs. IL difference</i>
Reinstatement	1,916 (481)	2,061 (571)	2,192 (671)	–131
Spillover	1,858 (489)	1,997 (599)	2,025 (473)	–28

Reinstatement sentence

The mean reading times for the reinstatement and spillover sentences for Experiment 1 appear in Table 1. There was a significant main effect of antecedent overlap condition, $F(2, 54) = 14.20$, $MSE = 40,220$, partial $\eta^2 = .35$; $F(2, 42) = 6.93$, $MSE = 82,258$, partial $\eta^2 = .75$. Planned comparisons demonstrated that reading times were shorter in the correct condition than in the incorrect–high-overlap condition, $F(1, 27) = 8.39$, $MSE = 74,951$, $d = 0.27$; $F(1, 21) = 5.55$, $MSE = 82,015$, $d = 0.42$. Reading times in the correct condition were also shorter than those in the incorrect–low-overlap condition, $F(1, 27) = 20.08$, $MSE = 113,635$, $d = 0.47$; $F(1, 21) = 9.28$, $MSE = 244,804$, $d = 0.37$. More important, reading times were shorter in the incorrect–high-overlap condition than in the incorrect–low-overlap condition, $F(1, 27) = 9.78$, $MSE = 52,692$, $d = 0.21$; $F(1, 21) = 4.15$, $MSE = 166,731$, $d = 0.71$.

Spillover sentence

The main effect of antecedent overlap was significant, $F(2, 54) = 4.85$, $MSE = 49,266$, partial $\eta^2 = .15$; $F(2, 42) = 4.04$, $MSE = 46,926$, partial $\eta^2 = .16$. Reading times were shorter in the correct condition than in the incorrect–high-overlap condition (although only marginal when based on items variability), $F(1, 27) = 4.56$, $MSE = 125,539$, $d = 0.25$; $F(1, 21) = 3.32$, $p = .08$, $MSE = 89,585$, $d = 0.32$. Reading times were also shorter in the correct condition than in the incorrect–low-overlap condition, $F(1, 27) = 13.95$, $MSE = 60,002$, $d = 0.35$; $F(1, 21) = 12.42$, $MSE = 59,579$, $d = 0.58$. Most important, the difference between the incorrect–high-overlap and incorrect–low-overlap conditions was not significant, both F s < 1, d s < .18.

The results of Experiment 1 are not consistent with the predictions made by the full-resolution account. Instead, they are consistent with the goodness-of-fit account, in which initial integration of information is based on its goodness of fit with the contents of working memory. When readers first encountered and integrated the anaphor in the reinstatement sentence, time to process this information varied as a function of the overlap between the anaphor and the antecedent; reading times were shorter in the correct condition than in the incorrect–high-overlap condition and shorter in those two conditions than in the incorrect–low-overlap condition. Readers continued to process the anaphor after moving past it, however. Reading times on the spillover sentence were

slow in both incorrect–overlap conditions, relative to the correct condition. This is inconsistent with an account that assumes that readers fully resolve the anaphor when they first encounter it. Instead, the results of the spillover sentence seem more consistent with a view in which initial integration of information is followed by a subsequent process in which information is verified or validated against additional information in memory. Further discussion of this view will be postponed until after the presentation of Experiments 2 and 3.

Experiment 2

In the previous experiment, the antecedent was always backgrounded by several sentences of text before the anaphor was encountered in the reinstatement sentence. Previous researchers have found that increased distance between an anaphor and its antecedent can lead to increased time to activate and reinstate the antecedent upon encountering the anaphor (O'Brien, 1987; O'Brien et al., 1995; O'Brien et al., 1990; O'Brien et al., 1997), but the effects of distance manipulations on anomalous anaphors have not been previously investigated. The materials from Experiment 1 were modified to create two antecedent distance conditions: antecedent near and antecedent distant. In the antecedent-near condition, the background section was omitted, and the reinstatement sentence directly followed the elaboration section, whereas in the antecedent-distant condition, the background was presented between the elaboration section and the reinstatement sentence. All passages were modified to maintain local coherence across all conditions.

If the goodness-of-fit effects observed on the reinstatement sentence in Experiment 1 were due to a failure to fully reactivate and reinstate the antecedent (Klin et al., 2006; Klin et al., 2004; Levine et al., 2000), those effects should be reduced when the distance between the anaphor and the antecedent is reduced (see also Rayner, Chace, Slattery, & Ashby, 2006). That is, in the antecedent-near conditions, readers may be more likely to fully reactivate and reinstate the antecedent; any difference in reading times between the incorrect–high- and –low-overlap conditions on the reinstatement sentence should be reduced or eliminated, and both conditions should yield longer reading times than the correct condition; reading times on the spillover sentence should not vary as a function of condition. Reading times in the antecedent-distant conditions should replicate those from the

reinstatement and spillover sentences in Experiment 1. On the other hand, pronoun researchers working with relatively short texts (e.g., Greene et al., 1992; Stewart et al., 2007) have suggested that distance between a referent and its antecedent may not be a driving force in whether readers will fully resolve an anaphor or not. If this is true, the pattern of effects should replicate those observed in Experiment 1, regardless of whether the antecedent is near or distant from the anaphor. Reading times on the reinstatement sentence should be shorter in the correct condition than in the incorrect–high-overlap condition, and both conditions should yield shorter reading times than the incorrect–low-overlap condition. On the spillover sentence, reading times should be shorter in the correct condition than in the two incorrect–overlap conditions, and there should be no difference between the two incorrect–overlap conditions.

Method

Participants and design

Seventy-two University of New Hampshire undergraduates enrolled in introductory psychology courses participated in exchange for partial course credit.

Two within-subjects variables were manipulated in this study: antecedent condition and antecedent distance. As in the previous experiment, antecedent condition had three levels: correct, incorrect–high-overlap, and incorrect–low-overlap. Antecedent distance had two levels: antecedent near and antecedent distant. The data for the reinstatement and spillover sentences were analyzed separately.

Materials

The 24 passages from Experiment 1, in addition to 6 comparable passages, were altered for use in this experiment (see the second example in the Appendix). Modifications consisted of minor rewording of the elaboration and background sections to ensure that local coherence was maintained in all conditions. The modified elaboration sections were, on average, 82.73, 82.63, and 83.20 words for the correct, incorrect–high-overlap, and incorrect–low-overlap conditions,

respectively. The modified background section (mean = 73.9 words) was omitted in the antecedent–near conditions. All other portions of the passages were the same as in Experiment 1.

Ten filler passages were included to balance the number of passages that did/did not contain anomalous information. As in the previous experiments, filler passages were similar in structure to the experimental passages, and they were followed by an equal number of “yes” and “no” questions. Six materials sets were constructed and counterbalanced such that an equal number of passages appeared in each condition, and across the materials sets, each passage appeared once in each condition.

Procedure

The procedure was the same as in Experiment 1.

Results and discussion

Comprehension accuracy rates were quite high (over 90 %). The reading times for Experiment 2 are presented in Table 2. There were no differences in comprehension accuracy as a function of condition, $F_s < 1$.

Reinstatement sentence

As in the previous experiment, there was a significant main effect of antecedent overlap, $F_1(2, 132) = 24.58, MSE = 100,247, \text{partial } \eta^2 = .27; F_2(2, 48) = 16.89, MSE = 59,754, \text{partial } \eta^2 = .41$. The effect of antecedent distance was unreliable, $F_1(1, 66) = 2.9, MSE = 63,400, p = .09, \text{partial } \eta^2 = .04; F_2 < 1, \text{partial } \eta^2 = .02$. More important, the antecedent overlap \times distance interaction was not significant, $F_s < 1, \text{partial } \eta^2_s < .03$. Planned comparisons revealed that the pattern of effects observed in Experiment 1 was replicated in both the antecedent–near and –distant conditions in the present study; thus, the two distance conditions will be reported together. The correct condition yielded shorter reading times than did the incorrect–high-overlap condition [antecedent near, $F_1(1, 66) = 14.12, MSE = 99,357, d = 0.33; F_2(1, 24) = 16.97, MSE = 32,433, d = 0.4$; antecedent distant (although not significant when based on items variability), $F_1(1, 66) = 5.86, MSE = 102,226, d = .23; F_2(1, 24) = 2.77, MSE = 104,138, p = .11, d = 0.24$]. The correct condition also yielded shorter

Table 2 Mean reading times (and standard deviations) in milliseconds for the reinstatement and spillover sentences as a function of antecedent overlap and distance conditions in Experiment 2

		Correct	Incorrect–High-Overlap (IH)	Incorrect–Low-Overlap (IL)	IH vs. IL difference
Antecedent distant	Reinstatement	1,769 (376)	1,860 (427)	2,009 (471)	–149
	Spillover	1,720 (343)	1,802 (399)	1,822 (392)	–20
Antecedent near	Reinstatement	1,780 (380)	1,919 (450)	2,062 (547)	–143
	Spillover	1,717 (371)	1,784 (378)	1,810 (380)	–26

reading times than did the incorrect–low-overlap condition [antecedent near, $F_1(1, 66) = 32.72$, $MSE = 174,900$, $d = 0.6$; $F_2(1, 24) = 36.55$, $MSE = 68,635$, $d = 0.72$; antecedent distant, $F_1(1, 66) = 27.94$, $MSE = 148,629$, $d = 0.56$; $F_2(1, 24) = 15.41$, $MSE = 101,899$, $d = 0.52$]. Finally, the incorrect–high-overlap condition yielded shorter reading times than did the incorrect–low-overlap condition [antecedent near, $F_1(1, 66) = 11.17$, $MSE = 130,726$, $d = 0.29$; $F_2(1, 24) = 7.53$, $MSE = 94,131$, $d = 0.35$; antecedent distant, $F_1(1, 66) = 11.25$, $MSE = 141,923$, $d = 0.33$; $F_2(1, 24) = 6.05$, $MSE = 84,759$, $d = 0.31$].

Spillover sentence

The pattern of effects from the spillover sentence also replicated those observed in the previous experiment. The main effect of antecedent overlap condition was significant, $F_1(2, 132) = 12.28$, $MSE = 30,764$, partial $\eta^2 = .17$; $F_2(2, 48) = 11.54$, $MSE = 13,161$, partial $\eta^2 = .33$. Neither the main effect of antecedent distance nor the overlap \times distance interaction approached significance, all $F_s < 1$, partial $\eta^2 < .01$. As with the reinstatement sentence reading times, the results for the antecedent-near and antecedent-distant conditions will be presented together. Planned comparisons demonstrated that reading times in the correct condition were shorter than those in the incorrect–high-overlap condition [antecedent near, $F_1(1, 66) = 4.63$, $MSE = 70,007$, $d = 0.18$; $F_2(1, 24) = 8.11$, $MSE = 29,122$, $d = 0.29$; antecedent distant, $F_1(1, 66) = 6.03$, $MSE = 80,800$, $d = 0.22$; $F_2(1, 24) = 3.85$, $MSE = 32,215$, $d = 0.25$]. The correct condition also yielded shorter reading times than did the incorrect–low-overlap condition [antecedent near, $F_1(1, 66) = 12.88$, $MSE = 58,824$, $d = 0.25$; $F_2(1, 24) = 7.16$, $MSE = 29,108$, $d = 0.34$; antecedent distant, $F_1(1, 66) = 16.35$, $MSE = 38,176$, $d = 0.28$; $F_2(1, 24) = 6.45$, $MSE = 52,592$, $d = 0.31$]. Finally, there was no difference in reading times between the two incorrect–overlap conditions in either the antecedent-near or -distant condition, all $F_s < 1$, $d_s < 0.08$.

As in Experiment 1, initial processing of the anaphor was a function of the degree of overlap between the anaphor and the antecedent. In addition, even after readers had moved on past the reinstatement sentence, readers continued to experience processing difficulty in the incorrect–overlap conditions; spillover sentence reading times were long in both incorrect–overlap conditions, relative to the correct condition. The combined pattern of effects from the reinstatement and spillover sentences in Experiments 1 and 2 are inconsistent with a full-resolution account and, instead, are consistent with predictions made by a goodness-of-fit account.

Furthermore, the pattern of effects observed here was not influenced by the degree of distance between the anaphor and the antecedent, despite previous demonstrations of distance effects in the antecedent retrieval literature (O'Brien, 1987; O'Brien et al., 1995; O'Brien et al., 1990; O'Brien et al., 1997).

However, focusing on antecedent characteristics in the elaboration may have exaggerated the role that conceptual information played in integration of the anaphor, regardless of distance. This issue will be explored in the next experiment.

Experiment 3

Results from the previous experiments demonstrated that the effects of the goodness of fit between the anaphor and the antecedent on initial processing of the anaphor are robust. They occurred even when the anaphor was unambiguous, there were no distractor antecedents in the text, and the distance between the antecedent and the anaphor was minimal. It is clear that readers did not fully activate and integrate the specific lexical item representing the antecedent immediately upon encountering the anaphor; if they had, no differences should have been observed between the two incorrect–overlap conditions. Instead, it appears that initial processing of the anaphor was based on the overlap between the anaphor and the contents of working memory. The specific lexical item was not fully integrated until after the anaphor had been processed; the difference between the two incorrect–overlap conditions was eventually eliminated, but not until readers had already moved on to the spillover sentence.

The question that is left unanswered, then, is the following: What information are readers reactivating and integrating when they first encounter the anaphor? Klin et al. (2006) offered two possible explanations. First, they suggested that upon reading the anaphor, readers may have only activated conceptual information about the antecedent (i.e., basic semantic information about the antecedent from general world knowledge), such that they could easily integrate that information with the anaphor without activating the specific lexical item. Alternatively, Klin et al. (2006) suggested that the readers may have activated sufficient information from the previously encountered passage context surrounding the antecedent, such that they were able to easily integrate the anaphor without activating the specific lexical item. The present experiment was designed to further explore these accounts.

The results of the previous experiments are, on the surface, consistent with the first account offered by Klin et al. (2006) and inconsistent with the second. Reading times on the reinstatement sentence increased as the conceptual overlap between the anaphor and the antecedent decreased. If readers had been using reactivated contextual information (i.e., information explicitly stated in the text) to integrate the anaphor, as suggested by the second account, they should have reactivated the distinguishing characteristics of the antecedent, which would have been difficult to integrate in both incorrect conditions. A third possibility is that readers initially reactivated contextual information, but because the retrieval mechanism responsible for antecedent retrieval (i.e., resonance; Myers &

O'Brien, 1998; O'Brien & Myers, 1999) is assumed to be both cyclical and continuous, this may have led to the activation of additional related information from general world knowledge. The elaboration sections in the previous experiments focused on distinguishing characteristics of the antecedent. According to this third account, activated contextual information concerning distinguishing characteristics of the antecedent may have led to the activation of other (nondistinguishing) characteristics of the antecedent from readers' general world knowledge. If all of this conceptual information became available quickly, regardless of its source, any of it could have been used to integrate the anaphor. That is, focusing on conceptual characteristics in the elaboration may have exaggerated the role that antecedent characteristics in general (regardless of source) played during initial integration of the anaphor, as compared with any influence of the specific lexical item representing the antecedent. This is consistent with arguments made by O'Brien and Albrecht (1991), who found that reactivating contextual information about the antecedent can lead to quick reactivation and reinstatement of related information from general world knowledge (see also Cook et al., 2001), even if the true antecedent was explicitly stated in the text.

In this experiment, the elaboration sections were rewritten such that no features of the antecedent were described (see the third example in the Appendix). However, the number of mentions of the antecedent (i.e., one explicit and two implicit) were the same as in the first two experiments. If, upon encountering the anaphor in the reinstatement sentence, readers only initially activated and integrated conceptual information about the antecedent (e.g., characteristics), regardless of context, the pattern of results should be the same here as in the previous experiments. Reading times on the reinstatement sentence should be shorter in the incorrect–high-overlap condition than in the incorrect–low-overlap condition, and both should yield longer reading times than in the correct condition. Alternatively, reactivating contextual information would not yield any characteristics of the anaphor; if initial processing of the anaphor is based on the ease of linking the anaphor with currently activated information, regardless of whether that information is conceptual or contextual in nature, the effects observed in the present experiment should differ from those in the previous experiments. There should be no difference in reinstatement sentence reading times between the two incorrect–overlap conditions, although both should yield longer reading times than the correct condition.

As was mentioned previously, the antecedent elaboration sections in the previous experiments described characteristics of the antecedent, which could have strengthened the role that conceptual information played in integration of the anaphor, regardless of distance. O'Brien et al. (1990) found that elaboration effects “trumped” distance effects in antecedent retrieval; elaborated distant antecedents were reactivated faster than

unelaborated near antecedents. Thus, in the present experiment, when conceptual information about the antecedent is unelaborated, distance effects may be more likely to emerge.

Method

Participants and design

Sixty University of Massachusetts undergraduates participated in exchange for course credit.

Two within-subjects variables were manipulated in this study: antecedent condition and antecedent distance. As in the previous experiment, antecedent condition had three levels: correct, incorrect–high-overlap, and incorrect–low-overlap. Antecedent distance had two levels: antecedent near and antecedent distant. The data for the reinstatement and spillover sentences were analyzed separately.

Materials

The 30 passages from Experiment 2 were modified for use in this experiment. First, the elaboration sections were rewritten so that no features of the antecedent were mentioned. As in the previous experiments, each antecedent was mentioned once explicitly and twice implicitly. The only difference between the three antecedent conditions was the antecedent itself; otherwise, the three versions were identical. The mean length of the elaboration section for all passages was 80.33 words. Second, as a result of changes to the elaboration sections, some of the background sections had to be rewritten to maintain local coherence. The rewritten background sections had a mean length of 75.27 words. All other aspects of the passages were the same as in the previous experiments.

Twenty filler passages were included to ensure that there was an equal number of anomalous and nonanomalous passages. There were an equal number of “yes” and “no” questions for the filler passages. Six materials sets were constructed and counterbalanced such that an equal number of passages appeared in each condition, and across the materials sets, each passage appeared once in each condition.

Procedure

The procedure was the same as that used in Experiments 1 and 2.

Results and discussion

The reading times for Experiment 3 are presented in Table 3. As in the previous experiments, there was no effect of passage condition on comprehension question accuracy rates (all p s > .80), F s < 1.

Table 3 Mean reading times (and standard deviations) in milliseconds for the reinstatement and spillover sentences as a function of antecedent overlap and distance conditions in Experiment 3

		Correct	Incorrect–High-Overlap (IH)	Incorrect–Low-Overlap (IL)	IH vs. IL difference
Antecedent distant	Reinstatement	1,827 (490)	1,964 (588)	2,030 (632)	–66
	Spillover	1,720 (445)	1,808 (413)	1,873 (433)	–65
Antecedent near	Reinstatement	1,790 (493)	1,974 (509)	2,024 (578)	–50
	Spillover	1,687 (429)	1,799 (403)	1,830 (433)	–31

Reinstatement sentence

As in Experiments 1 and 2, there was a significant main effect of antecedent overlap for the reinstatement sentence, $F_1(2, 108) = 20.69$, $MSE = 74,077$, partial $\eta^2 = .28$; $F_2(2, 48) = 11.93$, $MSE = 69,754$, partial $\eta^2 = .33$. Neither the main effect of distance nor the overlap \times distance interaction approached significance, all $F_s < 1$, $d_s < .01$. Because the patterns of effects were similar for the antecedent-near and antecedent-distant conditions, they will be presented together. Planned comparisons revealed that reading times were shorter for the correct antecedent condition than for the incorrect–high-overlap condition [antecedent near, $F_1(1, 59) = 17.59$, $MSE = 115,236$, $d = 0.37$; $F_2(1, 29) = 9.42$, $MSE = 102,834$, $d = 0.47$; antecedent distant, $F_1(1, 59) = 9.66$, $MSE = 117,055$, $d = 0.25$, $F_2(1, 29) = 8.69$, $MSE = 127,388$, $d = 0.51$]. Reading times were also shorter in the correct condition than in the incorrect–low-overlap condition [antecedent near, $F_1(1, 59) = 22.82$, $MSE = 143,268$, $d = 0.43$; $F_2(1, 29) = 12.42$, $MSE = 150,408$, $d = 0.6$; antecedent distant, $F_1(1, 59) = 12.26$, $MSE = 201,199$, $d = 0.36$; $F_2(1, 29) = 10.35$, $MSE = 101,171$, $d = 0.52$]. In contrast to the patterns observed in Experiments 1 and 2, there was no reliable difference between the two incorrect–overlap conditions, [antecedent near, $F_1(1, 59) = 1.19$, $MSE = 123,959$, $p > .27$, $d = 0.09$; $F_2(1, 29) = 2.28$, $MSE = 64,191$, $p > .14$, $d = 0.18$; antecedent distant, $F_1(1, 59) = 2.27$, $MSE = 113,408$, $p > .13$, $d = 0.11$; $F_2 < 1$, $d = 0.01$].

Spillover sentence

The main effect of antecedent condition was significant, $F_1(2, 108) = 16.65$, $MSE = 41,048$, partial $\eta^2 = .24$; $F_2(2, 48) = 11$, $MSE = 39,252$, partial $\eta^2 = .31$, but the main effect of distance was not, $F_1(1, 54) = 2.64$, $MSE = 27,6734$, $p = .11$, partial $\eta^2 = .05$; $F_2(1, 24) = 1.52$, $MSE = 21,864$, $p = .23$, partial $\eta^2 = .06$, nor was the antecedent overlap \times distance interaction, $F_s < 1$. Again, the patterns of effects were similar for the antecedent-near and antecedent-distant conditions, so they will be presented together. Planned comparisons revealed that reading times in the correct antecedent condition were shorter than those in the incorrect–high-overlap condition [antecedent near (although only when based on participants variability), $F_1(1, 59) = 8.21$, $MSE = 91,170$, $d = 0.27$; $F_2(1,$

$29) = 2.73$, $MSE = 63,662$, $p = .11$, $d = 0.29$; antecedent distant (although only marginal when based on items variability), $F_1(1, 59) = 5.54$, $MSE = 62,485$, $d = 0.21$; $F_2(1, 29) = 4.07$, $MSE = 69,814$, $p = .05$, $d = 0.39$]. The correct condition also yielded shorter reading times than did the incorrect–low-overlap condition, [antecedent near, $F_1(1, 59) = 19.5$, $MSE = 63,063$, $d = 0.33$; $F_2(1, 29) = 13.41$, $MSE = 60,118$, $d = .52$; antecedent distant, $F_1(1, 59) = 17.78$, $MSE = 78,904$, $d = 0.35$, $F_2(1, 29) = 11.44$, $MSE = 87,834$, $d = 0.61$]. Finally, the difference between the two incorrect conditions was not reliable [antecedent near, $F_1(1, 59) = 1.2$, $MSE = 49,713$, $p > .27$, $d = 0.08$; $F_2(1, 29) = 3.3$, $MSE = 70,237$, $p = .08$, $d = 0.31$; antecedent distant (only reliable when based on participants variability), $F_1(1, 59) = 4.07$, $MSE = 63,553$, $d = 0.16$; $F_2(1, 29) = 2.3$, $MSE = 95,720$, $p = .14$, $d = 0.29$].

It appears that the goodness-of-fit effects on the reinstatement sentence in the previous experiments were based on the ease with which the anaphor could be integrated with other activated information in working memory. In those experiments, that “other” information consisted primarily of characteristics of the antecedent, reactivated from both the previously read passage context and general world knowledge; as the number of characteristics that the anaphor and antecedent had in common increased, reading time for the reinstatement sentence decreased. However, in the present study, when contextual information did not describe any characteristics of the antecedent, readers may have been more likely to activate and integrate the specific lexical item representing the antecedent more quickly upon encountering the anaphor. Under these circumstances, the anaphor did not “fit” well with either of the incorrect–overlap antecedents, and reading time on the reinstatement sentence was lengthened in both conditions, relative to the correct condition.

General discussion

The present study investigated processing of anomalous anaphors that were unambiguous and for which there was only one possible antecedent. Previous research on antecedent retrieval has indicated that antecedents are quickly reactivated and reinstated under such conditions (e.g., O'Brien, 1987; O'Brien et al., 1997), suggesting that readers may fully resolve

anaphors immediately upon encountering them. The results of Experiments 1 and 2 were not consistent with a full-resolution account. Instead, initial integration of the anaphor appeared to be driven by the goodness of fit between the anaphor and the contents of active memory. Correct anaphors yielded shorter processing times on the reinstatement sentence than did incorrect but highly related anaphors, and both yielded shorter processing times than did incorrect but low-related anaphors.

The purpose of Experiment 3 was to explore whether the initial goodness-of-fit effects observed on the reinstatement sentence in Experiments 1 and 2 were based on the fit between the anaphor and the conceptual information about the antecedent, between the anaphor and activated contextual information, or both. The antecedent elaborations in the first two experiments focused on the characteristics of the antecedent, and this may have, in turn, reactivated additional characteristics from general world knowledge; together, this may have boosted the role of conceptual overlap between the anaphor and antecedent in the initial integration process. In the third experiment, the context did not describe any characteristics of the antecedent, and the context was exactly the same across conditions, except for the antecedent itself; readers may have activated the specific antecedent more quickly than when the context focused on characteristics. In this case, integration of the anaphor was difficult in both incorrect overlap conditions, relative to the correct condition. Taken together, the reinstatement sentence reading time results of Experiments 1–3 support a view in which the initial integration of the anaphor is based on its goodness of fit with activated information in memory, regardless of whether that activated information derives from the text itself or general world knowledge.

The second major finding reported here was that processing does not end with integration of the anaphor. The results from the spillover sentence in Experiments 1–3 are consistent with a view in which the linkages formed during the initial integration process are subsequently validated against information in long-term memory (Singer, 2013). Across all three experiments, both incorrect anaphors resulted in continued processing difficulty even after the reinstatement sentence had been processed. If the anaphor had been fully resolved when it was encountered in the reinstatement sentence, no such effects should have been observed on the spillover sentence. Instead, readers appeared to initially link the anaphor to the contents of active memory upon encountering it; the linkage formed during the first stage was then subsequently validated against information that may have become available after the integration process had begun.

The findings just described are consistent with assumptions of the RI-Val view, recently proposed by Cook and O'Brien (2014). This view assumes that reading involves three parallel asynchronous stages of processing: resonance (R), integration (I), and validation (Val). This view is an extension of previous two-stage activation+integration accounts of comprehension

(e.g., Cook & Myers, 2004; Kintsch, 1988; Long & Lea, 2005; Sanford & Garrod, 1989, 1998, 2005). In the first stage (R), information is reactivated from memory via a passive, dumb, and unrestricted retrieval mechanism, such as resonance (Myers & O'Brien, 1998; O'Brien & Myers, 1999). Information must be activated above some minimum threshold to impact subsequent stages of comprehension. Upon exceeding that threshold, activated information is then linked with the contents of working memory in the second, integration stage (I). Ease of forming this linkage may be based on the goodness of fit between the newly encountered information and the contents already in active memory. Similarly, as soon as some minimum level of integration has been completed, the validation (Val) of those linkages begins. Validation occurs against both the cumulative discourse model, which contains both previously encountered information from the text (such as the antecedent), as well as information from general world knowledge (Singer, 2013). The RI-Val view assumes that each stage is dependent upon the output of the preceding stage, and each stage is assumed to run to completion. Within the context of the present study, when the anaphor was encountered, it resulted in the reactivation (R) of related information in memory. As soon as activated content exceeded some minimum threshold, initial integration (I) of the anaphor with this activated content began; the ease of forming the linkage between the anaphor and the reactivated content (i.e., reading of the reinstatement sentence) was based on the goodness of fit. When readers subsequently attempted to validate (Val) that link against the discourse model, the discrepancy between the anaphor and the antecedent became apparent in the incorrect conditions, and subsequent processing (i.e., reading of the spillover sentence) was disrupted.

Although the goal of this study was to test the predictions of the goodness-of-fit account, it may be possible to explain more general shallow processing effects within the context of the RI-Val view. As was noted previously, many of the effects in the shallow processing literature are based on explicit anomaly detection tasks, which force a conscious binary choice on the reader and do not examine the time course of anomaly detection effects. Cook and O'Brien (2014) assumed that readers' standards of coherence (van den Broek et al., 1995) influence the extent to which readers will wait for all three processes (i.e., resonance, integration, and validation) to run to completion. O'Brien and Cook (*in press*) noted that if the output of the resonance and integration stages is sufficient to meet readers' standard of coherence, they may move on before validation is completed. That is, without full validation, shallow processing could occur, resulting in the failure to detect semantic anomalies (Barton & Sanford, 1993; Bohan & Sanford, 2008; Daneman et al., 2006; Erickson & Mattson, 1981; Hannon & Daneman, 2004; Sanford, 2002; Sanford et al., 2011).

The argument made here that processes involved in reading are both cyclical and continuous highlights the danger of oversimplifying predictions made by two-stage models prevalent in discourse comprehension studies (e.g., Gerrig & O'Brien, 2005; Glucksberg & McCloskey, 1981; Kintsch, 1988, 1998; Long & Lea, 2005; Rizzella & O'Brien, 2002; Sanford & Garrod, 1989, 1998, 2005). For explanatory purposes, these models tend to separate activation and integration into distinct stages. This can lead to the assumption that the reader moving on in the text signals that integration is complete. However, most of these models also contain the assumption that activation and integration stages are continuous and overlapping. In order to observe consequences of later-occurring activation and integration cycles, it is necessary to incorporate measures that allow for observation of the time course of processing across a larger window. This could include use of spillover regions, such as those used here, or use of eye-tracking technology (see Rayner, Pollatsek, Ashby, & Clifton, 2012). It is also important to incorporate assumptions about how and when integration mechanisms decide when tentative linkages made are sufficient for comprehension to proceed; this may require including an unconscious, but evaluative, component in the integration process (Long & Lea, 2005). The addition of the validation stage in the RI-Val model offers one potential solution.

The findings reported here also contribute to the ongoing discussion in the literature about the roles of semantic and contextual information in discourse comprehension. For example, Sanford and Garrod (1981, 1989, 2005; see also Garrod & Terras, 2000; Sanford, Garrod, Lucas, & Henderson, 1983) have argued for distinct bonding and resolution processes in discourse comprehension; bonding involves checking incoming information for its semantic fit with automatically reactivated information, whereas resolution involves verifying it against the actual discourse. Others, however, have argued that context has a strong initial influence on discourse interpretation (e.g., Hess, Foss, & Carroll, 1995; Nieuwland & Van Berkum, 2006; Otten & Van Berkum, 2007, 2008; Van Berkum, 2008). The results of the present study demonstrate that early stages of anaphoric processing can be dominated by either semantic or contextual information, depending upon which type of information, and how much, is reactivated first via a passive retrieval mechanism. However, additional information may become activated in subsequent activation cycles and, thus, influence later stages of processing. This is consistent with arguments made by Cook and colleagues (Cook & Guéraud, 2005; Cook & Myers, 2004; Cook & O'Brien, 2014; O'Brien & Cook, *in press*).

One final point is that although O'Brien and colleagues (O'Brien, 1987; O'Brien et al., 1990, 1995) have repeatedly found that reducing the distance between an anaphor and its antecedent results in reduced time to reactive and reinstate the

antecedent, no such effects were observed in Experiment 2 or 3. This may be because O'Brien and colleagues embedded anaphoric phrases in explicit demand sentences (e.g., *Mark's neighbor asked him what he had just finished building*) that would force coherence breaks if unresolved. In contrast, O'Brien et al. (1997) found that when a nominal anaphor was conceptually identical to its antecedent, no distance effects were observed. The anaphors used in the present study were much closer in nature to those used in the O'Brien et al. (1997) materials.

In conclusion, the present study extends findings from the shallow processing and anaphor-processing literatures by demonstrating that initial processing of unambiguous but anomalous anaphors appears to be based on their goodness of fit with the contents of active memory. However, this does not represent the end-stage of anaphoric processing; readers then validate those initial linkages against information in memory. These results are consistent with assumptions of the RI-Val view, proposed by Cook and O'Brien (2014).

Author Note Portions of this research were conducted in partial fulfillment of the requirements for the PhD at the University of New Hampshire. It was supported in part by a University of New Hampshire Dissertation Year Fellowship awarded to Anne E. Cook, National Science Foundation Grant SBR-9631040 awarded to Edward J. O'Brien, and by a Postdoctoral Traineeship to Anne E. Cook from NIMH training grant MH16745 awarded to the University of Massachusetts - Amherst. Portions of this research were presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL. The complete set of materials is available on request. I would like to thank Ed O'Brien, Jerry Myers, and Keith Rayner for their comments on a previous version of the manuscript.

Appendix

Sample passage from Experiment 1

Introduction

Terry loved classical music. She spent most of her waking hours listening to it, either in her room or in the car.

Correct Antecedent

Terry decided that it would be fun to teach herself how to play. She drove to a music shop located in the next town. As she entered the store she saw a beautiful cello. The large instrument was almost bigger than she was. Terry decided she wanted to learn how to play it. She imagined herself sitting down to play the heavy instrument. Terry asked the salesman for a price. After thinking for a few minutes, she decided to buy it that afternoon.

Incorrect–High-Overlap Condition

Terry decided that it would be fun to teach herself how to play. She drove to a music shop located in the next town. As she entered the store she saw a beautiful violin. It was very lightweight and fit perfectly between her chin and shoulder. She imagined herself dancing as she played beautiful music. Terry asked the salesman for a price. After thinking for a few minutes, Terry decided to buy it that afternoon.

Incorrect–Low-Overlap Condition

Terry decided that it would be fun to teach herself to play. She drove to a music shop located in the next town. As she entered the store she saw a beautiful oboe. The keys were bright and shiny, and the case was lined in black velvet. Terry decided she wanted to learn how to play it. She imagined herself fingering the keys to create perfect notes. Terry asked the salesman for a price. After thinking for a few minutes, she decided to buy it that afternoon.

Background

When Terry arrived home she found a message on her answering machine from her friend Jill. Because Terry hadn't spoken to Jill in over a week, she decided to invite her over for coffee. When Jill came over, she told Terry that she had a new boyfriend. After chatting about Jill's new boyfriend for a while, Jill asked Terry about what was new with her.

Reinstatement and Spillover Sentences

Terry showed her the cello she bought.
She even tried to play a few notes.

Closing

Terry told Jill that she was going to start practicing that very evening.

Sample passage from Experiment 2

Introduction

Terry loved classical music. She spent most of her waking hours listening to it, either in her room or in the car.

Correct Antecedent Condition

Terry decided to teach herself how to play. She and her friend Jill drove to a music shop located in the next town. As they entered the store, Terry saw a beautiful **cello**. The large instrument was almost bigger than she was. Terry decided she

wanted to learn how to play it. She imagined herself sitting down to play the heavy instrument. After thinking for a few minutes, she decided to buy it. Just then, Jill walked over to where Terry was standing.

Incorrect–High-Overlap Condition

Terry decided to teach herself how to play. She and her friend Jill drove to a music shop located in the next town. As they entered the store, Terry saw a beautiful violin. The small instrument fit perfectly between her chin and shoulder. Terry decided she wanted to learn how to play it. She imagined herself dancing as she played the lightweight instrument. After thinking for a few minutes, she decided to buy it. Just then, Jill walked over to where Terry was standing.

Incorrect–Low-Overlap Condition

Terry decided to teach herself how to play. She and her friend Jill drove to a music shop located in the next town. As they entered the store, Terry saw a beautiful oboe. The keys were bright and shiny, and the case was lined in black velvet. Terry decided she wanted to learn how to play it. She imagined herself fingering the keys to create perfect notes. After thinking for a few minutes, she decided to buy it. Just then, Jill walked over to where Terry was standing.

Background

Jill told Terry she wanted to stop at the craft store on the way home. She had been making a bunch of candle holders and had run out of supplies. She ran down her list and told Terry that she needed some paint, glue, glitter, and brushes. She also wanted to look to see if they had any stencils. After Jill finished her shopping, they went home and unloaded all of their purchases.

Reinstatement and Spillover Sentences

Terry showed Jill the cello she bought.
She even tried to play a few notes.

Closing

Terry told Jill that she was going to start practicing that very evening.

Sample passage from Experiment 3

Introduction

Terry loved classical music. She spent most of her waking hours listening to it, either in her room or in the car.

Correct Antecedent Condition

Terry decided that it would be fun to teach herself how to play. She and her friend Jill drove to a music shop located in the next town. As they entered the store, Terry saw a beautiful cello. Her father had played one with the city's philharmonic orchestra when he was a young man. Knowing this made Terry really want to learn how to play. After thinking for a few minutes, she decided to buy it. Just then, Jill walked over to where Terry was standing.

Incorrect–High-Overlap Condition

Terry decided that it would be fun to teach herself how to play. She and her friend Jill drove to a music shop located in the next town. As they entered the store, Terry saw a beautiful violin. Her father had played one with the city's philharmonic orchestra when he was a young man. Knowing this made Terry really want to learn how to play. After thinking for a few minutes, she decided to buy it. Just then, Jill walked over to where Terry was standing.

Incorrect–Low-Overlap Condition

Terry decided that it would be fun to teach herself how to play. She and her friend Jill drove to a music shop located in the next town. As they entered the store, Terry saw a beautiful oboe. Her father had played one with the city's philharmonic orchestra when he was a young man. Knowing this made Terry really want to learn how to play. After thinking for a few minutes, she decided to buy it. Just then, Jill walked over to where Terry was standing.

Background

Jill told Terry she wanted to stop at the craft store on the way home. She had been making a bunch of candle holders and had run out of supplies. She ran down her list and told Terry that she needed some paint, glue, glitter, and brushes. She also wanted to look to see if they had any stencils. After Jill finished her shopping, they went home and unloaded all of their purchases.

Reinstatement and Spillover Sentences

Terry showed her the cello she bought.
She even tried to play a few notes.

Closing

Terry told Jill that she was going to start practicing that very evening.

References

- Barton, S. B., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition*, *21*(4), 477–487.
- Bohan, J., & Sanford, A. (2008). Semantic anomalies at the borderline of consciousness: An eye-tracking investigation. *The Quarterly Journal of Experimental Psychology*, *61*(2), 232–239.
- Bredart, S., & Docquier, M. (1989). The Moses illusion: A follow-up on the focalization effect. *European Bulletin of Cognitive Psychology*, *9*(3), 357–362.
- Bredart, S., & Modolo, K. (1988). Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica*, *67*, 135–144.
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*(4), 368–407.
- Cook, A. E., & Guéraud, S. (2005). What have we been missing? The role of general world knowledge in discourse processing. *Discourse Processes*, *39*(2–3), 265–278.
- Cook, A. E., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language*, *44*(2), 220–234.
- Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory and Language*, *50*(3), 268–288.
- Cook, A. E., Myers, J. L., & O'Brien, E. J. (2005). Processing an anaphor when there is no antecedent. *Discourse Processes*, *39*(1), 101–120.
- Cook, A. E., & O'Brien, E. J. (in press). Passive activation and instantiation of inferences during reading. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading*. Cambridge University Press.
- Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*, 26–49.
- Daneman, M., Hannon, B., & Burton, C. (2006). Are there age-related differences in shallow semantic processing of text? Evidence from eye movements. *Discourse Processes*, *42*(2), 177–203.
- Duffy, S. A., & Rayner, K. (1990). Eye movements and anaphor resolution: Effects of antecedent typicality and distance. *Language and Speech*, *33*(2), 103–119.
- Ehrlich, S. F., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, *22*, 75–87.
- Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, *20*, 540–551.
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.
- Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*(1–2), 71–83.
- Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language*, *42*, 526–544.
- Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, *39*(2–3), 225–242.
- Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(5), 311.

- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371.
- Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 266.
- Hannon, B., & Daneman, M. (2004). Shallow semantic processing of text: An individual-differences account. *Discourse Processes*, *37*(3), 187–204.
- Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, *124*(1), 62.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354.
- Kamas, E. N., Reder, L. M., & Ayers, M. S. (1996). Partial matching in the Moses illusion: Response bias not sensitivity. *Memory & Cognition*, *24*(6), 687–699.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Klin, C. M., Guzmán, A. E., Weingartner, K. M., & Ralano, A. S. (2006). When anaphor resolution fails: Partial encoding of anaphoric inferences. *Journal of Memory and Language*, *54*(1), 131–143.
- Klin, C. M., Weingartner, K. M., Guzmán, A. E., & Levine, W. H. (2004). Readers' sensitivity to linguistic cues in narratives: How salience influences anaphor resolution. *Memory & Cognition*, *32*(3), 511–522.
- Levine, W. H., Guzmán, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *Journal of Memory and Language*, *43*(4), 594–617.
- Long, D. L., & Lea, R. B. (2005). Have we been searching for meaning in all the wrong places? Defining the "search after meaning" principle in comprehension. *Discourse Processes*, *39*(2–3), 279–298.
- Love, J., & McKoon, G. (2011). Rules of engagement: Incomplete and complete pronoun resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 874.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2&3), 131–157.
- Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111.
- Nieuwland, M. S., & Van Berkum, J. J. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, *106*(2), 119–131.
- O'Brien, E. J., Albrecht, J. E., Hakala, C. M., & Rizzella, M. L. (1995). Activation and suppression of antecedents during reinstatement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 626–634.
- O'Brien, E. J., & Cook, A. E. (in press). Models of discourse comprehension. In A. Pollatsek & R. Treiman (Eds.), *Handbook on reading*. Oxford University Press.
- O'Brien, E. J. (1987). Antecedent search processes and the structure of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 278–290.
- O'Brien, E. J., & Albrecht, J. E. (1991). The role of context in accessing antecedents in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 94–102.
- O'Brien, E. J., & Myers, J. L. (1999). Text comprehension: A view from the bottom up. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 35–53). Mahwah, NJ: Lawrence Erlbaum Associates.
- O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 241–249.
- O'Brien, E. J., Raney, G. E., Albrecht, J. E., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes*, *23*(1), 1–24.
- Otten, M., & Van Berkum, J. J. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research*, *1153*, 166–177.
- Otten, M., & Van Berkum, J. J. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*(6), 464–496.
- Poesio, M., Sturt, P., Artstein, R., & Filik, R. (2006). Underspecification and anaphora: Theoretical issues and preliminary evidence. *Discourse Processes*, *42*(2), 157–175.
- Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241–255.
- Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. (2012). *Psychology of reading* (2nd ed.). New York, NY: Psychology Press.
- Reder, L. M., & Cleeremans, A. (1990). The role of partial matches in comprehension: The Moses illusion revisited. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation* (Vol. 25, pp. 233–258). San Diego, CA: Academic Press, Inc.
- Reder, L. M., & Kusbit, G. W. (1991). Locus of the Moses illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language*, *30*, 385–406.
- Rizzella, M. L., & O'Brien, E. J. (2002). Retrieval of concepts in script-based texts and narratives: The influence of general world knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 780–790.
- Sanford, A. J. (2002). Context, attention and depth of processing during interpretation. *Mind & Language*, *17*(1–2), 188–206.
- Sanford, A. J., & Emmott, C. (2012). *Mind, brain and narrative*. Cambridge University Press.
- Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language: Explorations of comprehension beyond the sentence*. New York: Wiley.
- Sanford, A. J., & Garrod, S. C. (1989). What, when, and how?: Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes*, *4*, 235–262.
- Sanford, A. J., & Garrod, S. C. (1998). The role of scenario mapping in text comprehension. *Discourse Processes*, *26*(2–3), 159–190.
- Sanford, A. J., & Garrod, S. C. (2005). Memory-based approaches and beyond. *Discourse Processes*, *39*(2–3), 205–224.
- Sanford, A. J., Garrod, S., Lucas, A., & Henderson, R. (1983). Pronouns without explicit antecedents? *Journal of Semantics*, *2*(3–4), 303–318.
- Sanford, A. J., & Graesser, A. C. (2006). Shallow processing and underspecification. *Discourse Processes*, *42*(2), 99–108.
- Sanford, A. J., Leuthold, H., Bohan, J., & Sanford, A. J. (2011). Anomalies at the borderline of awareness: An ERP study. *Journal of Cognitive Neuroscience*, *23*(3), 514–523.
- Sanford, A. J., Sanford, A. J., Filik, R., & Molle, J. (2005). Depth of lexical-semantic processing and sentential load. *Journal of Memory and Language*, *53*(3), 378–396.
- Sanford, A. J., Sanford, A. J., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, *42*(2), 109–130.
- Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences*, *6*(9), 382–386.
- Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science*, *22*(5), 361–366.

- Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language; Journal of Memory and Language*.
- Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence for delay. *The Quarterly Journal of Experimental Psychology*, 60(12), 1680–1696.
- Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review*, 11(5), 882–888.
- Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, 36(1), 201–216.
- Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39(4), 558–592.
- Van Berkum, J. J. (2008). Understanding sentences in context what brain waves can tell us. *Current Directions in Psychological Science*, 17(6), 376–380.
- Van den Broek, P., Risdien, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ward, P., & Sturt, P. (2007). Linguistic focus and memory: An eye movement study. *Memory & Cognition*, 35(1), 73–86.