

# Journal of Educational Psychology

## **The Language of Instruction: Compensating for Challenge in Lectures**

Srdan Medimorec, Philip I. Pavlik, Jr., Andrew Olney, Arthur C. Graesser, and Evan F. Risko  
Online First Publication, February 9, 2015. <http://dx.doi.org/10.1037/edu0000024>

### CITATION

Medimorec, S., Pavlik, P. I., Jr., Olney, A., Graesser, A. C., & Risko, E. F. (2015, February 9). The Language of Instruction: Compensating for Challenge in Lectures. *Journal of Educational Psychology*. Advance online publication. <http://dx.doi.org/10.1037/edu0000024>

# The Language of Instruction: Compensating for Challenge in Lectures

Srdan Medimorec  
University of Waterloo

Philip I. Pavlik, Jr., Andrew Olney, and  
Arthur C. Graesser  
University of Memphis

Evan F. Risko  
University of Waterloo

Recent studies have used Coh-Metrix, an automated text analyzer, to assess differences in language characteristics across different genres and academic disciplines (Graesser, McNamara, & Kulikowich, 2011; McNamara, Graesser, McCarthy, & Cai, 2014). Coh-Metrix analyzes text on many constructs at different levels, including Word Concreteness (vs. abstractness), Narrativity (vs. informational), Deep Cohesion, Referential Cohesion, and Syntactic Simplicity. In previous research, texts in the natural sciences had lower Narrativity and Word Concreteness than texts in the language arts, but were higher in Syntactic Simplicity and Referential Cohesion. This pattern suggests a form of compensation in which difficulty on one dimension (e.g., Word Concreteness) is compensated for by increasing text ease on another dimension (e.g., Syntactic Simplicity). In the present study, we provide a further test of this compensation idea by analyzing oral language use across humanities and natural science lectures. We demonstrate that decreases in Word Concreteness across lectures are associated with increases in Narrativity, Deep Cohesion, and Syntactic Simplicity. In addition, within lectures, decreases in Word Concreteness are associated with increases in Syntactic Simplicity. Compensatory mechanisms are discussed in this article at different levels of language and discourse.

*Keywords:* language, education, lectures, linguistic compensation, Coh-Metrix

Individuals vary different features of their speech and writing as a function of different situational characteristics (Ashburn & Gordon, 1981; Fraser & Roberts, 1975; Magnifico, 2010; McNamara, 2013; Walker, Risko, & Kingstone, 2014). For example, individuals produce more fillers (*um* and *uh*) when speaking to a human than a computer (Walker et al., 2014), caregivers adopt syntactically simpler language when talking to the elderly (Ashburn & Gordon, 1981), and speakers produce longer utterances in the absence of time pressure (Müller, Großmann-Hutter, Jameson, Rummer, & Wittig, 2001). Similarly, authors' choices of text features (e.g., sentence complexity, text cohesion) are driven by their readers' needs (Magnifico, 2010; McNamara, 2013). One potential function of such variation, at least from the speaker's and writer's perspective, could be to facilitate listener/reader comprehension (McNamara, 2013; Thornton & Light, 2006). As such, this behavior might be particularly important in instructional contexts. Indeed, recent work has uncovered evidence that linguistic characteristics of texts vary across different genres and academic

disciplines in a manner that suggests that writers are compensating for challenges on one (or more) dimensions of text difficulty by simplifying other dimensions (Graesser et al., 2014; Graesser, McNamara, & Kulikowich, 2011; McNamara, 2013; McNamara, Graesser, & Louwerse, 2012). For example, science texts normally have difficult conceptual topics, but they tend to have simpler syntax and higher referential cohesion than narrative texts (e.g., stories). In the current study, we expand on the investigation of these putative compensation mechanisms by exploring variation in speech characteristics during academic lectures.

## Text Measures Across Genres and Academic Disciplines

A series of recent studies (e.g., Graesser et al., 2011, 2014; Lightman, McCarthy, Dufty, & McNamara, 2007; McNamara, 2013; McNamara et al., 2012) have used Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara, Graesser, McCarthy, & Cai, 2014) to assess how features of texts differ across different genres and academic disciplines. Coh-Metrix is an automated text analyzer that computes various measures of cohesion and readability indices of text. It provides measures of five principal components (Word Concreteness, Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity) that are composed of 53 language-discourse measures. These top-five principal components accounted for over 54% of text variability among 37,520 paragraph-length texts in a multidimensional space (Graesser et al., 2011). Each of these components provides a measure of "ease" (or difficulty), in which higher scores are

---

Srdan Medimorec, Department of Psychology, University of Waterloo; Philip I. Pavlik, Jr., Andrew Olney, Arthur C. Graesser, Department of Psychology, University of Memphis; Evan F. Risko, Department of Psychology, University of Waterloo.

Correspondence concerning this article should be addressed to Srdan Medimorec, University of Waterloo, Department of Psychology, PAS Building, 200 University Avenue West, Waterloo, ON, Canada, N2L 3G1. E-mail: smedimor@uwaterloo.ca

associated with “easier” text on a given dimension. It is important to note that the labels for the principal components represent hypotheses about the underlying construct and should be interpreted with the possibility that there may be alternative interpretations. As a result, it is important to critically evaluate patterns at the level of the individual discourse measures that make up each component when interpreting results. In the present investigation, we focus on the principal components but also consider discourse measures independent of those components.

The Word Concreteness principal component is jointly determined by the measures of word concreteness, word meaningfulness, and word imageability (the values are derived from the Medical Research Council Psycholinguistic database; Coltheart, 1981). Sadoski, Kealy, Goetz, & Paivio (1997) identified a similar principal component in earlier work. Previous research has demonstrated that comprehension and memory for text and oral discourse increase with increases in these measures (Allen & Hulme, 2006; Bourassa & Besner, 1994; Clark & Paivio, 1991; Danks, 1969; Graesser et al., 2011; Hidi & Baird, 1988; Kroll & Merves, 1986; McNamara, 2013; Mosenthal, 1996; D. L. Nelson & Schreiber, 1992; Paivio, 1990; Sadoski & Paivio, 2013; Schraw, Bruning, & Svoboda, 1995; Strain, Patterson, & Seidenberg, 1995; Wade & Adams, 1990). Such effects have been explained in the context of dual coding theory (DCT; Clark & Paivio, 1991; Paivio, 1971; Sadoski & Paivio, 2013), in which the activation of two distinct mental codes—the verbal code and the nonverbal code—facilitates processing.

Narrativity refers to the extent to which a text can be categorized as narrative as opposed to informational (e.g., Biber, 1988). A high Narrativity score indicates a more story-like discourse (i.e., the frequent use of verbs, adverbs, and intentional actions, events, and particles). Another fundamental aspect of this component is that it reflects measures characteristic of everyday conversation, such as word familiarity and simpler syntactic constructions (e.g., the use of shorter noun phrases and fewer words before the main verb or clause). Previous studies (e.g., Graesser et al., 2011, 2014; McNamara et al., 2012) have demonstrated that actual narrative texts (as coded by other researchers) are consistently higher on this dimension than informational texts (e.g., science and social studies). Moreover, narrative texts are generally easier to comprehend, are read faster, and are recalled better than informational texts (Graesser, Hauff-Smith, Cohen, & Pyles, 1980; Haberlandt & Graesser, 1985). Narrativity scores negatively correlate with grade-level estimates very robustly (Graesser et al., 2011; J. Nelson, Perfetti, Liben, & Liben, 2012; Pearson & Hiebert, 2010).

High Deep Cohesion indicates frequent use of causal (e.g., because, so), temporal (e.g., after, before), logical (e.g., and then, if-then), additive (e.g., moreover, however), and/or adversative connectives (e.g., but, although)—a feature that suggests that ideas in text are explicitly linked (Cain & Nash, 2011; Crismore, Markkanen, & Steffensen, 1993; McNamara et al., 2014). When the links among the ideas in text are lacking, comprehension can be disrupted (e.g., Britton & Gulgoz, 1991; Zwaan & Radvansky, 1998).

High Referential Cohesion indicates frequent overlap (i.e., coreference) of nouns, arguments, stems, and content words between sentences, again implying linking of words and concepts between sentences (McNamara & Kintsch, 1996). Previous studies (e.g., Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975;

Vidal-Abarca, Martínez, & Gilabert, 2000) have demonstrated that increasing referential links increases text comprehension and recall.

Syntactic Simplicity is determined by several measures, such as the number of words per sentence and the number of words before the main verb of the main clause (i.e., simpler sentences will be shorter and contain fewer words before the main verb). Longer sentences increase comprehension difficulty because they place a greater load on working memory compared with shorter sentences (Caplan & Waters, 1999; but also see MacDonald, 1997).

In the following, when referring to Coh-Metrix principal components, we use capital letters in order to distinguish these references from references to related constructs (e.g., the measure word concreteness is a subcomponent of the Word Concreteness principal component). Nonetheless, each Coh-Metrix principal component represents a hypothesized measure of its associated construct (for more details, see Graesser et al., 2004, 2011, 2014; Li, Cai, & Graesser, 2013; McNamara et al., 2014).

Using Coh-Metrix, previous studies (Graesser et al., 2011, 2014; McNamara, 2013) reported that texts in the natural sciences were characterized by lower Narrativity and Word Concreteness than texts in the language arts, but higher in Syntactic Simplicity and Referential Cohesion. The authors suggested that this pattern reflected a linguistic trade-off in which the use of a large proportion of unfamiliar ideas (i.e., lower Narrativity) and abstract words (i.e., lower Word Concreteness) in the science texts was compensated for by the use of simpler syntax (e.g., shorter and less complex sentences) and more cohesion (e.g., greater word overlap). Thus, topic was argued to lead to the use of less narrative and less concrete language, and, in order to compensate for this difficulty, writers were held to use simpler syntax and more cohesive language. These findings are in line with other analyses of the linguistic features of narrative and instructional texts (McNamara et al., 2012, 2014).

### Why Investigate Linguistic Compensation?

The understanding of linguistic compensation has important implications for educational practice. For example, as a result of this putative linguistic compensation behavior, texts would rarely be “challenging” on all indices of text ease, a feature that, in theory, could aid comprehension (McNamara, 2013). In other words, linguistic compensation provides a potential route through which the “overall” difficulty of to-be-learned material can be controlled. Consistent with this idea, Beck and colleagues (Beck, McKeown, Sinatra, & Loxterman, 1991; Beck, McKeown, & Worthy, 1995) demonstrated that “abstract” textbook passages can be modified to enhance comprehension and recall by adding more action and conversational language (which would correspond to Narrativity in Coh-Metrix), cohesion (corresponding to Referential Cohesion in Coh-Metrix), and connectivity (corresponding to Deep Cohesion in Coh-Metrix), as well as breaking complex sentences into several simpler sentences (corresponding to Syntactic Simplicity in Coh-Metrix). Thus, comprehension and recall of more abstract texts can be enhanced by modifying other discourse features (e.g., adding more cohesion, simpler syntax).

Linguistic compensation, in practice, could be implemented to either simplify complex material or introduce challenge. As such, linguistic compensation can be viewed as a means of keeping

instructional discourse within an optimal zone of comprehension difficulty (i.e., not too easy, not too hard, e.g., Graesser et al., 2014; McNamara et al., 2014; Wolfe et al., 1998). A good illustration of the potential educational benefit of keeping text within an optimal zone of difficulty is the *reverse cohesion effect* (O'Reilly & McNamara, 2007), in which students with low knowledge benefit from added cohesion, whereas students with high knowledge benefit from less cohesive texts on measures that tap deeper comprehension. The lower cohesion encourages high-knowledge readers to be more active comprehenders and generate more inferences. In a similar vein, the expertise reversal effect, in which the optimal amount of challenge increases as expertise increases (Kalyuga, Ayres, Chandler, & Sweller, 2003), provides empirical support for the general notion of keeping educational material within an optimal zone. Critically, as this research demonstrates, “optimal” does not always mean “easier.” In the learning and cognitive sciences, this idea has been formalized as a trade-off between assistance (e.g., scaffolding) versus generative student activity when a student is learning, with the implication that there is a zone in which assistance is optimal for learning (Koedinger & Aleven, 2007; Koedinger, Pavlik, McLaren, & Aleven, 2008). Understanding the nature of this optimal zone of comprehension difficulty represents an important goal in educational psychology, and investigating how instructors manage this delicate balance to stay within this optimal zone (or fail to), whether in the written or spoken word, represents a step in this larger goal.

### Present Investigation

In this study, we further examine the notion of linguistic compensation. Specifically, we examine the idea that decreases in ease on Word Concreteness (a principal component made up of word concreteness, word meaningfulness, and word imageability) and Narrativity are associated with increases in ease on other dimensions (e.g., Deep Cohesion, Referential Cohesion, Syntactic Simplicity) in an instructional context. We expand on previous research by shifting from texts to instructional oral discourse, specifically, lectures. Research using Coh-Metrix has primarily focused on written text, but has also been used to analyze naturalistic oral discourse (Graesser, Jeon, & Cai, 2008; Graesser, Jeon, Yan, & Cai, 2007; Louwerse, McCarthy, McNamara, & Graesser, 2004; McNamara et al., 2014). Moreover, to the extent that both oral discourse and printed text compositions are intended to communicate informative content to an audience, we can expect them to have similar characteristics, namely, to be understandable and coherent (Graesser & McNamara, 2011). In addition, lectures involve only one voice, they are planned (roughly) in advance, and they are designed for a general audience rather than a single person (Stein, 1992; Tannen, 1982). Lectures are also the dominant teaching method in higher education (Bligh, 1972; Finkelstein, Seal, & Schuster, 1998; Lammers & Murphy, 2002; Pritchard, 2010). For example, Griffin and Cashin (1989) reported that almost 75% of college courses use lectures as a primary instructional method, and it does not appear that much has changed since that time (Lammers & Murphy, 2002; Pritchard, 2010; Watts & Schaur, 2011). Thus, understanding methods (e.g., linguistic compensation) of keeping material within an optimal zone of difficulty in lectures is important from an educational practice perspective. By understanding

language use in lectures, we can provide new insights into the processes involved in planning and delivering a lecture, on the one hand, and comprehension and retention of lecture content, on the other. If the kind of linguistic trade-off observed in texts represents a general feature of language use, then we should expect to find similar patterns in lectures (Graesser et al., 2011; McNamara, 2013; McNamara et al., 2012).

Although they share a number of characteristics, lectures and texts can differ in important ways. For example, Glass, Hazen, Hetherington, and Wang (2004) reported that language used in lectures was less organized and less concise than the corresponding textbook material, presumably a consequence of more material being composed extemporaneously instead of being preplanned. Oral language (e.g., lecture) is also less formal compared with texts (Graesser et al., 2014; Li et al., 2013). Texts afford more opportunity for the author to plan, edit, and purposefully craft the word choice, sentence construction, and other features of the message being communicated. Consequently, in the lecture context, it might be more difficult to engage in compensating linguistic behavior that apparently underlies the patterns of text difficulty previously observed across academic disciplines. On the other hand, in the “live” setting of a traditional lecture, speakers have more paralinguistic cues available to drive linguistic compensation (such as backchannel head nods and confused looks), which might encourage compensatory efforts. In addition, the lecturers also have some general knowledge about the audience (e.g., the students’ background knowledge)—another factor that might facilitate compensation (e.g., Schober & Brennan, 2003).

In addition to expanding the analysis to oral instructional discourse, we conducted a series of more direct tests of the compensation account of variation across genre/academic discipline than have been previously provided. Specifically, previous research compared the Coh-Metrix principal components’ patterns across genres and academic disciplines. For example, McNamara (2013) reported that the natural science texts were lower in Word Concreteness and Narrativity relative to the language arts texts, while also being higher in Referential Cohesion and Syntactic Simplicity compared with the language arts texts. These patterns were interpreted as evidence of linguistic compensation, in which challenges arising from lower Word Concreteness and lower Narrativity in the science texts were offset by simpler syntax and higher Referential Cohesion. We provide the equivalent analyses but also directly assess the compensation account through a series of regression analyses and an analysis of changes in discourse within a lecture.

There is an implied causal order according to the compensation account of variation in discourse characteristics across genres and academic disciplines. Specifically, topic difficulty influences a given principal component (i.e., Word Concreteness, Narrativity), and variation on that component causes offsetting changes on other principal components (i.e., Referential Cohesion, Syntactic Simplicity). We directly tested this prediction by conducting a series of mediation analyses with the putative driver of the compensation as the mediator, academic discipline as the independent variable, and the principal component putatively being offset to afford the compensation as the dependent variable. A mediation analysis also allows us to assess the extent to which evidence for compensation can be found independent of comparisons across academic discipline (e.g., science vs. the language arts). In other words, is there

evidence for compensation after controlling for differences in language use patterns across discipline? This test is critical if the notion of compensation is to be considered a natural part of instructional discourse rather than a feature of how individuals in the humanities or natural sciences use language (or write, in the case of texts).

We next turned to an analysis of variation in the Coh-Matrix principal components *within* lectures. Previous research (Lightman et al., 2007) has reported that text difficulty, indexed by Flesch-Kincaid Grade Level (FKGL; Klare, 1974–1975), increased linearly across chapters in history and science textbooks, whereas cohesion remained unchanged. On the other hand, grade level decreased across narrative texts. Thus, measures of text difficulty can change within a text. To investigate potential changes in language components across lectures, we broke each lecture into a set of time-based intervals (i.e., 5-min bins). Again, if lecturers compensate for difficulty on one dimension by simplifying discourse on other dimensions, then within a lecture, offsetting patterns of change over time should be present. For example, if content becomes less concrete/meaningful/imageable over time (indexed by the Word Concreteness principal component), then this should be compensated for by an increase in another dimension (e.g., Syntactic Simplicity). Confirmation of this prediction would represent a particularly strong demonstration of compensatory linguistic behavior because it would provide evidence for the behavior *within* a lecturer rather than across lecturers (as in the previous analyses).

In a final series of analyses, we explored the use of two composite measures of discourse characteristics that consist of combining scores across all of the Coh-Matrix principal components. The first composite measure consists of averaging across all five principal components in order to provide a single index of “ease”: (Equation 1 = [Word Concreteness + Narrativity + Deep Cohesion + Referential Cohesion + Syntactic Simplicity]/5). This measure is admittedly coarse. For example, it weights each dimension of ease equally. Nevertheless, it provides a tool to explore overall difficulty, which is an important construct in an investigation of compensatory behavior. If individuals are compensating for challenge on one dimension by simplifying others, then a measure of overall difficulty provides a (rough) index of the success of that compensatory effort. For example, perfect compensation would be reflected by trade-offs at the level of individual principal components across different lectures in conjunction with equivalent overall difficulty. The second composite measure we explore is *formality* (Graesser et al., 2014). Formal language is considered to be precise, coherent, and articulate, with abstract content and complex syntax, whereas informal language is less cohesive, with concrete words and simple syntax (Graesser et al., 2014). Informational texts are expected to be higher on this formality dimension than narrative texts. This composite measure is calculated by the following formula: (Equation 2 = [Referential Cohesion + Deep Cohesion – Word Concreteness – Narrativity – Syntactic Simplicity]/5). It should be noted that the formality measure correlates higher with standard metrics of text difficulty than does the overall ease measure (Graesser et al., 2014). Together, these analyses provide one of the most comprehensive tests of linguistic compensation in an instructional context.

## Method

### Tool Description

The lecture transcripts were analyzed using the Coh-Matrix text analyzer (Graesser et al., 2004, 2011; McNamara et al., 2014). All reported variables were computationally extracted from text, and no human coding was used. It should be noted that Coh-Matrix integrates a set of readily available text analyzing tools (e.g., lexical databases, syntactic parsers, part-of-speech taggers, lexical coreference tools; for more details see McNamara et al., 2014, Chapter 3). In addition, a free online version of Coh-Matrix is available at <http://tool.cohmatrix.com/>.

We focused on five principal components signifying ease (Word Concreteness, Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity), three descriptive indices (number of words, number of sentences, and sentence length), and two composite measures (overall ease and formality). These different measures are of course correlated to some extent. Table 1 presents correlations among principal components across academic discipline. In addition to various standard language and discourse measures (e.g., word count, syntactic categories, lexical overlap measures), Coh-Matrix computes *z* scores and percentiles for each of the five major principal components. These measures inform us about how principal components of the analyzed text compare with the Touchstone Applied Science Associates (TASA) corpus (Laudauer, Foltz, & Laham, 1998). The TASA corpus includes texts from various genres and academic disciplines between kindergarten and college-ready, and it contains over 11 million words. The five major components were constrained to be orthogonal in the TASA corpus in the Graesser et al. (2011) analysis. This was not true here and typically would not be when Coh-Matrix is used to analyze “new” texts/discourse. Nevertheless, it is interesting that the orthogonality of principal components was largely preserved when the TASA-based principal component norms were used in a sample of hundreds of state-nominated texts available in the Common Core (see Graesser et al., 2014); the correlations ranged from

Table 1  
*Correlations Among Principal Components*

	1	2	3	4	5
All lectures					
1. Word Concreteness	—				
2. Narrativity	-.64**	—			
3. Deep Cohesion	-.60**	.50**	—		
4. Referential Cohesion	-.35*	.52**	.30*	—	
5. Syntactic Simplicity	-.66**	.30*	.41**	.09	—
Humanities					
1. Word Concreteness	—				
2. Narrativity	-.53**	—			
3. Deep Cohesion	-.24	.46*	—		
4. Referential Cohesion	-.29	.61**	.30	—	
5. Syntactic Simplicity	-.09	-.17	-.08	-.41*	—
Natural sciences					
1. Word Concreteness	—				
2. Narrativity	-.43*	—			
3. Deep Cohesion	-.55**	.23	—		
4. Referential Cohesion	.10	.13	-.05	—	
5. Syntactic Simplicity	-.68**	.22	.39*	-.08	—

\*  $p < .05$ . \*\*  $p < .001$ .

-0.28 to +0.28 when pairwise correlations between principal components were computed.

## Corpus

Our corpus consisted of transcripts of 94 lectures (delivered by 54 lecturers), taken from the websites of three U.S. universities: Yale University (Open Yale Courses, <http://oyc.yale.edu>), Massachusetts Institute of Technology (MIT; OpenCourseWare, <http://ocw.mit.edu>), and the University of Michigan (MICASE Corpus, <http://quod.lib.umich.edu/m/micase>). There were 52 Yale lecture transcripts (40 humanities and 12 natural sciences), 28 MIT lecture transcripts (all natural sciences), and 14 University of Michigan transcripts (six humanities and eight natural sciences). Of those, 46 were in the humanities and 48 were in the natural sciences. Yale and MIT lecturers were represented by two lectures, and the University of Michigan lecturers were represented by one lecture (this choice was based on lecture transcript availability). All lectures were delivered in a university setting in front of students (i.e., our analysis does not include lectures delivered online or on air).

We analyzed 40 min of each Yale and MIT lecture, and an average of 5,500 words of each lecture from the MICASE corpus. Because the MICASE corpus does not include lecture video and audio recordings, and the lecture transcripts do not contain time markers, we analyzed an average of 5,500 words per lecture, which roughly corresponds to 40 min of lecturing (based on the Yale and MIT lecture analyses). For the between-lecture/lecturer analyses, the entire text (40 min or 5,500 words) was submitted to Coh-Metrix (94 texts). For the within-lecture analysis, we divided the Yale and MIT lecture transcripts into eight 5-min time bins (intervals) and submitted them to Coh-Metrix (640 texts).

Our choice of the Yale, MIT, and University of Michigan lectures was based on their mostly monologic character (i.e., single speaker, limited dialogue and interaction with the audience). Even though some of the lecturers used notes while delivering lectures, the majority of the lecture content was not read to the audience (although they were used as guidance in some cases, we cannot comment on the level of detail in the notes). In the case of Yale and MIT lectures, this was determined from videos of the lectures; in the case of the MICASE transcripts, the inference was made based on the presence of fillers, repetitions, and false starts. The transcripts of the lectures are openly available under the Creative Commons or Fair Use licenses (for terms of use, see <http://oyc.yale.edu/terms>, <http://ocw.mit.edu/terms>, and <http://micase.elicorpora.info>). We removed paragraph breaks, annotated parts of speech, actions of the speaker, and parts of the transcripts containing dialogue with students from all transcripts. Additionally, we did not edit out symbols or equations. Finally, we edited the MICASE corpus by adding the missing full stops, correcting informal colloquial expressions (i.e., “wanna,” “gonna,” “cuz”), and deleting false starts (Yale and MIT transcripts were already “clean”). Thus, our corpus-cleaning rule was consistent across lectures (for Coh-Metrix corpus-cleaning recommendations, see McNamara et al., 2014).

For the purpose of classification of academic disciplines, we adopted a slightly modified nomenclature used by Schachter, Christenfeld, Ravina, and Bilous (1991). We classified biology, biomedical engineering, biological/health science, chemistry, ecology, geology, microbiology, physics, geophysics, mathematics,

computer science, and electrical engineering as the natural sciences. Although the latter two would likely not be considered natural sciences proper, Schachter et al. included mathematics in that group and, as such, we thought computer science and electrical engineering best fit in that classification. This group of topics might better be considered what has come to be known as the STEM disciplines (i.e., science, technology, engineering, and math). Art history, African American studies, American studies, anthropology, classics, history, literature, music, philosophy, religious studies, and women’s studies were classified as the humanities. Although Schachter et al. included political science as a social science, one political science lecture we used was “communication,” which fits better with the humanities discipline.

## Results

### Principal Component and Subcomponent Correlations

Our first set of analyses explored the correlations between each Coh-Metrix principal component and its subcomponents, and between the principal components themselves (Word Concreteness, Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity). For each lecturer from the Yale and MIT sample, in which two lectures from each lecturer were used, the scores were averaged across their two lecture samples (i.e., the individual lectures were submitted to Coh-Metrix and the scores were averaged to yield one score per lecturer). Thus, the analyses are based on 26 humanities and 28 natural sciences lectures (except when noted). Although the Coh-Metrix measures have been validated (for more details, see McNamara et al., 2014), the goal of our analysis was to explore whether principal components and their subcomponents correlated in the expected direction in our corpus. We expected to find positive correlations between principal components and their positively loaded subcomponents, and negative correlations between principal components and their negatively loaded subcomponents (see Graesser et al., 2011). As expected, the overwhelming majority of the analyzed subcomponents (41 out of 42) correlated in the expected direction with their respective principal components. All Word Concreteness (absolute range of correlations = .92 to .97), Deep Cohesion (.15 to .91), Referential Cohesion (.24 to .92), and Syntactic Simplicity (.66 to .91) subcomponents correlated with their respective principal components in the expected direction. One of the Narrativity subcomponents (i.e., third-person pronoun incidence) did not correlate in the expected direction with its principal component, but the remaining did (absolute range of correlations = .18 to .90).

A second set of correlational analyses focused on the relations between the principal components, both overall and within each academic discipline. It is important to clarify the meaning of the correlations between principal components. Originally, the five principal components were constrained to be orthogonal (Graesser et al., 2011), following the standard methodology of principal component analysis. However, orthogonality is not expected for a corpus of texts in a restricted age span, knowledge domain, or context. In particular, this study focuses on lectures for college students, so the correlations among principal components are informative and expected to deviate from orthogonality. The notion of linguistic compensation can therefore be meaningfully explored. As is clear in Table 1, overall, there are negative correlations

between Word Concreteness and the other principal components, and positive correlations between the other components. This general pattern held, with a few exceptions, within both the humanities and natural science lectures.

### The Effect of Academic Discipline

Our next set of analyses explored how descriptive indices of text and Coh-Matrix principal components differed across the humanities and the natural science lectures. A series of regression analyses were performed with academic discipline (0 = humanities vs. 1 = natural sciences) as the predictor, and descriptive indices of text (i.e., number of words, number of sentences, and average sentence length) and Coh-Matrix principal components (i.e., Word Concreteness, Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity) as the outcome, controlled for academic institution (the three universities, dummy coded). Means across academic disciplines for the descriptive indices of text and the Coh-Matrix principal components are presented in Table 2, and the differences between means are shown in Figure 1. In this section and throughout, 95% confidence intervals are provided in square brackets (i.e., [lower limit, upper limit]) and  $r_s$  is the semipartial correlation.

**Descriptive indices.** The regression analysis in which academic discipline was the predictor of the average number of words per lecture was based on Yale and MIT lectures only (because we analyzed 40 min of those lectures). There was no statistical difference in the average number of words per lecture across the humanities (5,622 [5301, 5943]) and the natural sciences (5,383 [5061, 5704]),  $B = 213 [-434, 860]$ ,  $SE = 319$ ,  $t(38) = .67$ ,  $p = .51$ ,  $r_s = .10$ . Using all lectures, there was a significant effect of academic discipline on the mean number of sentences per lecture,  $B = 49.67 [9.82, 89.52]$ ,  $SE = 19.84$ ,  $t(52) = 2.50$ ,  $p = .016$ ,  $r_s = .30$ , such that there were more sentences in the natural science lectures (324 [301, 346]) than in the humanities lectures (268 [245, 291]). In conjunction with the equal number of words, this translated into a significant effect of academic discipline on the mean number of words per sentence,  $B = -2.62 [-5.08, -.16]$ ,  $SE = 1.23$ ,  $t(52) = -2.14$ ,  $p = .037$ ,  $r_s = -.24$ , such that the sentences in the humanities lectures contained more words on average (21.6

Table 2

*Means and 95% CIs for the Descriptive Indices (Mean Values), Principal Components (Z Scores), and Composite Measures (Z Scores) Across the Humanities and the Natural Science Lectures*

Measure	Humanities M [95% CI]	Natural sciences M [95% CI]
Number of Words <sup>a</sup>	5622 [5301, 5943]	5383 [5061, 5704]
Number of Sentences	268 [245, 291]	324 [301, 346]
Words per Sentence	21.6 [20.1, 23.0]	17.4 [16.0, 18.8]
Word Concreteness	-.88 [-1.10, -.65]	-1.93 [-2.15, -1.70]
Narrativity	.39 [.26, .52]	.76 [.63, .89]
Deep Cohesion	.71 [.49, .93]	1.33 [1.11, 1.54]
Referential Cohesion	-.05 [-.20, .10]	.32 [.18, .47]
Syntactic Simplicity	-.13 [-.30, .03]	.44 [.28, .60]
Overall Ease	.01 [-.06, .07]	.19 [.12, .25]
Formality	.26 [.19, .33]	.48 [.41, .54]

Note. CI = confidence interval.

<sup>a</sup> Yale and Massachusetts Institute of Technology lectures only.

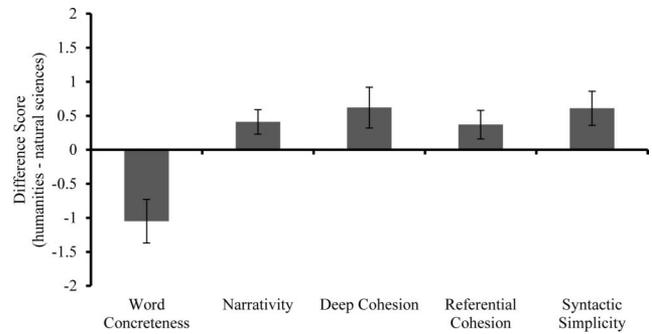


Figure 1. Differences between the principal components' mean  $z$  scores in the humanities and the natural sciences lectures (difference = humanities – natural sciences). Error bars indicate the 95% confidence intervals. Bars above the horizontal axis indicate that the natural sciences are “easier” than the humanities lectures, and bars below the horizontal axis indicate that humanities are “easier” than the natural sciences.

[20.1, 23.0]) compared with the sentences in the natural science lectures (17.4 [16.0, 18.8]). In summary, although there was no difference in the overall number of words per lecture between the humanities and the natural sciences, the humanities lectures contained longer sentences and fewer sentences compared with the natural science lectures.

**Principal components.** For all principal components, lower scores indicate more difficulty or complexity, whereas higher scores indicate more ease or simplicity, on a given measure. Academic discipline had a significant effect on Word Concreteness, such that it was higher in the humanities ( $-.88 [-1.10, -.65]$ ) than the natural sciences ( $-1.93 [-2.15, -1.70]$ ),  $B = -.74 [-1.09, -.38]$ ,  $SE = .18$ ,  $t(52) = -4.12$ ,  $p < .001$ ,  $r_s = -.37$ . The opposite pattern (i.e., natural sciences “easier” than humanities) was found in Narrativity (natural sciences, .76 [.63, .89]; humanities, .39 [.26, .52]),  $B = .26 [.02, .50]$ ,  $SE = .12$ ,  $t(52) = 2.18$ ,  $p = .034$ ,  $r_s = .26$ , (marginally) in Deep Cohesion (natural sciences, 1.33 [1.11, 1.54]; humanities, .71 [.49, .93]),  $B = .37 [-.02, .76]$ ,  $SE = .19$ ,  $t(52) = 1.91$ ,  $p = .061$ ,  $r_s = .23$ , (marginally) in Referential Cohesion (natural sciences, .32 [.18, .47]; humanities,  $-.05 [-.20, .10]$ ),  $B = .26 [-.01, .53]$ ,  $SE = .14$ ,  $t(52) = 1.90$ ,  $p = .063$ ,  $r_s = .24$ , and in Syntactic Simplicity (natural sciences, .44 [.28, .60]; humanities,  $-.13 [-.30, .03]$ ),  $B = .35 [.09, .62]$ ,  $SE = .13$ ,  $t(52) = 2.65$ ,  $p = .011$ ,  $r_s = .27$ . Thus, the natural science lectures were more abstract (i.e., less concrete, meaningful, and imageable), but somewhat more narrative-like, cohesive, and syntactically simpler, than the humanities lectures.<sup>1</sup>

### Mediation Analysis

Results of the preceding analysis revealed a somewhat different pattern than that found in previous work. Specifically, Narrativity was higher (not lower) in the natural sciences than in the humanities. We return to why this might be the case in the General

<sup>1</sup> Qualitatively similar patterns were observed when we compared humanities and natural sciences within Yale lectures and within MICASE lectures.

Discussion. For the present purposes, this is inconsistent with the notion that individuals are compensating for lower Narrativity in the natural sciences by modulating other language discourse dimensions. Thus, the following analyses will focus on the hypothesis that individuals are compensating for decreasing word concreteness, meaningfulness, and imageability (i.e., Word Concreteness driven linguistic compensation) in the natural sciences by simplifying discourse on the other dimensions (e.g., Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity). It is important to note that this causal order is theoretically motivated by the idea that abstract conceptualizations are considered more difficult (Allen & Hulme, 2006; Bourassa & Besner, 1994; Danks, 1969; Graesser et al., 2011; Hidi & Baird, 1988; Kroll & Merves, 1986; McNamara, 2013; Mosenthal, 1996; D. L. Nelson & Schreiber, 1992; Paivio, 1990; Strain et al., 1995; Wade & Adams, 1990). The General Discussion further addresses the plausibility of this causal order. Other potential models are of course possible, given the correlational nature of the design, so we also include analyses of alternative models.

According to the Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) driven compensation account, academic discipline should influence Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity indirectly, through Word Concreteness (the mediating variable). In a similar vein, the mediation analysis provides a test of the extent to which the Word Concreteness principal component is related to the other principal components, independent of academic discipline (i.e., the *b* path). A series of regressions were performed with academic discipline as the predictor, each principal component's *z* score as the outcome, and the Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) *z* score as the mediator, after controlling for academic institution (a dummy-coded covariate). We used the standard mediation annotation, in which *a* = the unstandardized regression coefficient between academic discipline and the mediator, *b* = the unstandardized regression coefficient between the mediator and the dependent variable, *c* = the unstandardized regression coefficient between academic discipline and the dependent variable, and *c'* = the unstandardized regression coefficient between academic discipline and the dependent variable controlling for the mediator. To estimate indirect effects and effect sizes for the indirect effects in our mediation models, we used the PROCESS modeling tool (Hayes, 2012), set to Model 4, with 1,000 bootstrap samples and a 95% confidence level for confidence intervals. We report the ratio of indirect to total effect ( $P_M$ ; e.g., Hayes, 2013; MacKinnon, 1994; MacKinnon & Dwyer, 1993; Preacher & Hayes, 2008; Preacher & Kelley, 2011; Sobel, 1982) as the estimate of the effect size of the indirect effect. Table 3 illustrates the results of the mediation analyses.

The results of the mediation analyses demonstrated that Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) mediated the relation between academic discipline and Narrativity, Deep Cohesion, and Syntactic Simplicity.<sup>2</sup> In other words, variation in Narrativity, Deep Cohesion, and Syntactic Simplicity across academic discipline can be accounted for by variation in Word Concreteness (i.e., word concreteness, word meaningfulness, and word imageability) across academic discipline, when controlling for academic institution. However, variation in Referential Cohesion across discipline could not be ac-

counted for by variation in the Word Concreteness principal component (see Table 3).

### Further Exploration of Mediation Models

The mediation models described in the Mediation Analysis section test our proposed causal order relating academic discipline to changes in Coh-Matrix principal components. However, given the correlational design, it is informative to consider both alternative models and different parameterizations of those models. In the following, we describe the alternative mediation models and the results of the analyses.

**Reverse causal models.** As noted our proposed model is that academic discipline is causing changes in word concreteness, word meaningfulness, and word imageability, which leads to changes in the other principal components. Although arguably more difficult to motivate theoretically, academic discipline could cause changes in the other principal components, which then lead to changes in Word Concreteness (i.e., reversing the mediator and the dependent variable in our original analysis). In our first set of alternative mediations, we explored this possibility (the predictor and the covariates remained the same as in our original analysis). Because Word Concreteness did not mediate the relation between academic discipline and Referential Cohesion, we considered only models including Word Concreteness, Narrativity, Deep Cohesion, and Syntactic Simplicity. In this set of analyses, Narrativity and Syntactic Simplicity produced significant indirect effects, though effect sizes in all models were similar. In addition, the relation between academic discipline and Word Concreteness (i.e., direct effect) remained statistically significant when the mediator was introduced in all the reverse mediation models (see Table 4).

**Mediation models covarying out the influence of other principal components.** As is clear in Table 1, the different principal components are related to one another. As such, a stronger case for mediation (in the case of our theoretical model) could be made if Word Concreteness mediated the relation between academic discipline and the other principal components while controlling for the remaining principal components (and academic institution). Thus, we repeated the previous analyses with academic discipline as the predictor, each principal component's *z* score as the outcome, and Word Concreteness *z* score as the mediator, controlled for academic institution (a dummy coded covariate) and the remaining principal components (i.e., those not functioning as the mediator or dependent variable, including referential cohesion). Similar to our original analyses, Word Concreteness mediated the relation between academic discipline and Narrativity, Deep Cohesion, and Syntactic Simplicity, even when controlling for the other principal components and academic institution (see Table 5).

We also conducted a similar analysis using the reverse causal models described in the Reverse Causal Models section. Here, the relation between academic discipline (the predictor) and Word Concreteness (the outcome) was mediated by the other individual

<sup>2</sup> When we performed a series of mediation analyses on Yale and MICASE lectures only, we observed qualitatively similar patterns. Moreover, Word Concreteness correlated negatively with the other principal components (with the exception of Referential Cohesion) in the Yale natural science lectures, MIT natural science lectures, and MICASE natural science lectures.

Table 3  
*Mediation Analyses: Word Concreteness as the Mediator of the Relation Between Academic Discipline (the Predictor) and the Other Principal Components (Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity), Controlled for Academic Institution (Dummy Coded)*

Outcome	Unstandardized regression coefficients				Indirect effect [95% CI]	Effect size
	a	b	c	c'		
Narrativity	-.74***	-.32***	.26**	.02	.23 [.09, .47]	.90
Deep Cohesion	-.74***	-.47**	.37*	.02	.35 [.13, .67]	.93
Referential Cohesion	-.74***	-.01	.26*	.25	.01 [-.21, .22]	.03
Syntactic Simplicity	-.74***	-.20*	.35**	.21	.15 [.02, .38]	.42

\*  $p < .10$ . \*\*  $p < .05$ . \*\*\*  $p < .001$ .

principal components (i.e., Narrativity, Deep Cohesion, and Syntactic Simplicity), controlled for the remaining principal components (other than the mediator and the outcome), Referential Cohesion, and academic institution (dummy coded). When controlling for principal components not functioning as the potential mediator or dependent variable in the model, only one of the reverse causal models was significant, specifically, Syntactic Simplicity (see Table 6).

Given the reverse causal model with Syntactic Simplicity successfully mediating the relation between academic discipline and Word Concreteness, while controlling for academic institution, Narrativity, Deep Cohesion, and Referential Cohesion, we wanted to assess Syntactic Simplicity (in the same manner as Word Concreteness) as a potential mediator with respect to the other Coh-Matrix principal components. In other words, does Syntactic Simplicity (like Word Concreteness) also mediate the relation between academic discipline and Narrativity and Deep Cohesion? Results suggest that Syntactic Simplicity did mediate the relation between academic discipline and Referential Cohesion (in this case, the effect of academic discipline on Referential Cohesion increased with the inclusion of the mediator), but did not appear to mediate the relation between academic discipline and Narrativity or Deep Cohesion. In contrast, the original mediation analysis shows that Word Concreteness mediates the relation between academic discipline and Syntactic Simplicity, Narrativity, and Deep Cohesion (see Table 7).

In summary, the test of the alternative mediation models demonstrated that Word Concreteness mediated the relation between academic discipline and Narrativity, Deep Cohesion, and Syntactic Simplicity, even when controlling for the other principal components and academic institution. Moreover, the only other plausible

reverse causal model when the other principal components were controlled was Syntactic Simplicity. However, Syntactic Simplicity mediated the relation between academic discipline and Word Concreteness, and academic discipline and Referential Cohesion, but not the relation between academic discipline and either Narrativity or Deep Cohesion.

### Principal Components' Subcomponent Analysis

Although the emphasis of the preceding analyses was on descriptive measures and principal components, it is critical to recognize that each principal component is made up of multiple discourse measures. Assessing patterns at the level of the subcomponents can provide both a more nuanced view of the putative compensatory patterns and a test of the principal components themselves. To this end, we conducted a set of partial correlational analyses (controlling for academic discipline) between Word Concreteness and the subcomponents of each principal component. We included only the measures with primary loadings on principal components (see Graesser et al., 2011). Because of the overall negative correlations between Word Concreteness and the other Coh-Matrix principal components, the subcomponents with a positive loading on a given principal component are expected to correlate negatively with Word Concreteness, whereas the subcomponents with a negative loading are expected to correlate positively with Word Concreteness.

The results of these analyses indicated that the majority of Narrativity, Syntactic Simplicity, and Deep Cohesion subcomponents correlated with Word Concreteness in the direction indicated by the overall correlations between Word Concreteness and the

Table 4  
*Relation Between Academic Discipline (the Predictor) and Word Concreteness Mediated by the Other Individual Principal Components (Narrativity, Deep Cohesion, and Syntactic Simplicity), Controlled for Academic Institution (Dummy Coded)*

Mediator	Unstandardized regression coefficients				Indirect effect [95% CI]	Effect size
	a	b	c	c'		
Narrativity	.26**	-.72***	-.74***	-.55**	-.18 [-.38, -.03]	.25
Deep Cohesion	.37*	-.40**	-.74***	-.59***	-.15 [-.41, .02]	.20
Syntactic Simplicity	.35**	-.36*	-.74***	-.61**	-.13 [-.32, -.01]	.17

\*  $p < .10$ . \*\*  $p < .05$ . \*\*\*  $p < .001$ .

Table 5

*Word Concreteness as the Mediator of the Relation Between Academic Discipline and the Other Principal Components (Narrativity, Deep Cohesion, and Syntactic Simplicity), Controlled for the Remaining Principal Components (Other Than the Mediator and the Outcome), Referential Cohesion, and Academic Institution (Dummy Coded)*

Outcome	Unstandardized regression coefficients				Indirect effect [95% CI]	Effect size
	a	b	c	c'		
Narrativity	-.47**	-.30**	.11	-.03	.14 [.03, .38]	1.31
Deep Cohesion	-.44***	-.38**	.15	-.02	.17 [.03, .46]	1.13
Syntactic Simplicity	-.51**	-.21*	.39**	.28*	.11 [.01, .34]	.28

\*  $p < .10$ . \*\*  $p < .05$ . \*\*\*  $p < .001$ .

principal components in Table 1.<sup>3</sup> Specifically, 14 of the 16 Narrativity subcomponents (see Table 8), all of the Deep Cohesion subcomponents (see Table 9), and all of the Syntactic Simplicity subcomponents (see Table 10) correlated in the expected direction. This was not the case for the Referential Cohesion subcomponents (see Table 11). Of the 12 Referential Cohesion subcomponents, five correlated with Word Concreteness in the expected direction (i.e., content word overlap–adjacent sentences, content word overlap – all sentences, type–token ratio [TTR; Templin, 1957], lexical diversity – all words, and latent semantic analysis [LSA] givenness versus newness), and seven in the opposite direction (i.e., argument overlap – adjacent sentences, argument overlap – all sentences, noun overlap – adjacent sentences, stem overlap – all sentences, lexical diversity – verbs, LSA overlap – adjacent sentences, and LSA overlap – all sentences). Thus, for Narrativity, Syntactic Simplicity, and Deep Cohesion, but not Referential Cohesion, the overall pattern with respect to the relation between Word Concreteness and each principal component is consistent across the respective subcomponents.

### Within-Lecture Analysis

In order to broaden our test of the compensation account of linguistic patterns in lectures, we next explored potential changes in language characteristics in the humanities and the natural science lectures as a function of time in the lecture. For this analysis, we divided the Yale and MIT lecture transcripts into eight 5-min time bins (intervals) and submitted them to Coh-Metrix (640 texts). The two lectures from each lecturer were averaged prior to analysis. For the mean values of all measures across intervals see Table 12.

**Descriptive measures of text.** For each descriptive, we calculated a slope relating time to the particular descriptive measure in question and compared the individual slope values with zero (using one-sample  $t$  tests). Thus, a positive score would indicate that particular measure was increasing over time, whereas a negative score would indicate that particular measure was decreasing over time. In addition, we compared these slope values across academic disciplines. The number of words per interval did not change significantly across time (i.e., across eight lecture intervals; .939 [-1.681, 3.559]),  $t(39) = .72$ ,  $SEM = 1.29$ ,  $p = .47$ ,  $d = .12$ . On the other hand, the number of sentences increased across time (.448 [.148, .747]),  $t(39) = 3.02$ ,  $SEM = .15$ ,  $p = .004$ ,  $d = .48$ , and the average sentence length per interval decreased across time (-.272 [-.408, -.137]),  $t(39) = -4.07$ ,  $SEM = .07$ ,  $p < .001$ ,

$d = .64$ . Finally, the slope relating time and average sentence length was steeper in the humanities (-.41 [-.59, -.23]) than the natural sciences (-.13 [-.32, .05]),  $F(1, 38) = 4.68$ ,  $MSE = .16$ ,  $p = .03$ ,  $d = .69$ . The slope values of the other descriptive measures of text did not differ significantly across the humanities and natural science lectures.

**Principal components.** Another series of analyses, using the same design, was performed in order to explore the relation between the principal components'  $z$  scores and the 5-min lecture intervals. The mean slope values for each principal component are presented in Figure 2. The results of the analysis revealed that Word Concreteness (-.022 [-.041, -.003]) decreased across time,  $t(39) = -2.33$ ,  $SEM = .009$ ,  $p = .025$ ,  $d = .37$ . Narrativity (-.005 [-.016, .006]) did not change significantly across time,  $t(39) = -.93$ ,  $SEM = .006$ ,  $p = .36$ ,  $d = .15$ . The same was true for Deep Cohesion (-.019 [-.045, .007]),  $t(39) = -1.47$ ,  $SEM = .013$ ,  $p = .40$ ,  $d = .23$ , and Referential Cohesion (-.015 [-.036, .006]),  $t(39) = -1.46$ ,  $SEM = .015$ ,  $p = .15$ ,  $d = .23$ . Lastly, Syntactic Simplicity (.033 [.014, .053]) increased across time,  $t(39) = 3.55$ ,  $SEM = .009$ ,  $p = .001$ ,  $d = .56$ .

Thus, whereas Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) decreased, Syntactic Simplicity increased as a function of time in the humanities and natural science lectures (see Figure 3 for the mean values of Word Concreteness and Syntactic Simplicity across lecture intervals). No other principal component changed significantly across time in the lecture.<sup>4</sup>

### Composite Measures

Our next set of analyses explored how two composite measures differed across the humanities and the natural science lectures and how they changed across lectures. We used the same design as the previous analyses. The two composite measures were overall ease (i.e., Equation 1) and formality (i.e., Equation 2).

**Effect of academic discipline.** Academic discipline had a significant effect on overall ease, such that it was higher in the

<sup>3</sup> Given the number of correlations conducted to support this analysis (and the resulting familywise error rate), we choose to focus qualitatively on direction of the correlation rather than on the significance or lack thereof of a given correlation.

<sup>4</sup> Flesch-Kincaid Reading Ease increased across lecture intervals (.33 [.05, .61]),  $t(39) = 2.40$ ,  $SEM = .005$ ,  $p = .021$ ,  $d = .38$ . Moreover, the Word Concreteness principal component correlated negatively with the Flesch-Kincaid Reading Ease,  $r(52) = -.71$ ,  $p < .001$ .

Table 6  
*Relation Between Academic Discipline (the Predictor) and Word Concreteness (the Outcome) Mediated by the Other Individual Principal Components (Narrativity, Deep Cohesion, and Syntactic Simplicity), Controlled for the Remaining Principal Components (Other Than the Mediator and the Outcome), Referential Cohesion, and Academic Institution (Dummy Coded)*

Mediator	Unstandardized regression coefficients				Indirect effect [95% CI]	Effect size
	a	b	c	c'		
Narrativity	.11	-.66**	-.47**	-.40**	-.07 [-.31, .08]	.15
Deep Cohesion	.15	-.25*	-.44**	-.40**	-.04 [-.22, .05]	.08
Syntactic Simplicity	.39**	-.29*	-.51***	-.40**	-.11 [-.33, -.02]	.22

\*  $p < .10$ . \*\*  $p < .05$ . \*\*\*  $p < .001$ .

natural sciences (.19 [.12, .25]) than humanities (.01 [-.06, .07]),  $B = .10$  [-.01, .22],  $SE = .06$ ,  $t = 1.79$ ,  $p < .001$ ,  $r_s = .21$ . In addition, academic discipline had a significant effect on formality, such that it was higher in the natural sciences (.48 [.41, .54]) than the humanities (.26 [.19, .33]),  $B = .15$  [.02, .28],  $SE = .06$ ,  $t = 2.41$ ,  $p = .019$ ,  $r_s = .28$ . Thus, the natural science lectures contained more formal language than the humanities lectures. As noted in the Present Investigation section, the overall ease is a coarse measure that correlates lower with standard metrics of text difficulty than does formality, which might explain a seemingly counterintuitive pattern reported here (i.e., natural sciences more formal, but easier than, humanities).

**Within-lecture analysis.** We calculated a slope relating time to each composite measure and compared that value with zero (Yale and MIT lectures only). Thus, a positive score would indicate that a measure was increasing over time, whereas a negative score would indicate that a measure was decreasing over time. Overall ease (-.006 [-.014, .002]) did not change significantly across lecture,  $t(39) = -1.41$ ,  $SEM = .004$ ,  $p = .17$ ,  $d = .22$ , and neither did formality (-.008 [-.018, .002]),  $t(39) = -1.60$ ,  $SEM = .005$ ,  $p = .12$ ,  $d = .25$ . There were no significant differences in the slope values across the humanities and natural science lectures. Thus, neither of the composite measures demonstrated any linear change over time within lecture.

### Using Individual Discourse Measures to Test for Linguistic Compensation

The present study explored linguistic compensation by relying on the Coh-Metrix principal components. This of course does not imply that linguistic compensation cannot be investigated outside of the Coh-Metrix principal component context. In the following

section, we demonstrate that similar analyses (with similar outcomes) can be conducted without the theoretical commitment to the Coh-Metrix principal components, by using individual linguistic measures that can indicate text difficulty/ease. Here, we conduct a series of regression and mediation analyses by using the same design as the previous analyses. The tables with the results are provided in the Table 13 (regression analyses) and Table 14 (mediation analyses).

For this set of analyses, we manually calculated the word concreteness composite measure by averaging word concreteness, word meaningfulness, and word imageability scores. Thus, all three subcomponents were equally weighted in this new measure (i.e., average word concreteness/meaningfulness/imageability), and this was the only composite measure used in the analysis. Next, we selected eight discourse measures that indicate text ease/difficulty (i.e., causal connectives incidence, content word overlap—all sentences, FKGL, Measure of Textual Lexical Diversity [MTLD, McCarthy & Jarvis, 2010], modifiers per noun phrase, noun incidence, passive voice incidence, and TTR). Although these measures are taken from the Coh-Metrix output, they can also be assessed independent of the Coh-Metrix tool.

Each of the individual measures we selected has previous work supporting its connection with difficulty/ease and, as such, represent useful targets for compensatory efforts. Previous studies have demonstrated that the use of causal connectives in text and oral discourse clarifies the relations between ideas and the lack of connectives hinders comprehension, at least for low knowledge readers (Britton & Gulgoz, 1991; McNamara, Crossley, & McCarthy, 2010; Zwaan & Radvansky, 1998). Similarly, increased content word overlap enhances text comprehension and recall (Beck et al., 1991; Kintsch, 1998). On the other hand, an increase in FKGL

Table 7  
*Syntactic Simplicity as the Mediator of the Relation Between Academic Discipline and the Other Principal Components (Narrativity, Deep Cohesion, and Referential Cohesion), Controlled for Academic Institution (Dummy Coded)*

Outcome	Unstandardized regression coefficients				Indirect effect [95% CI]	Effect size
	a	b	c	c'		
Narrativity	.35**	-.04	.26**	.27**	-.01 [-.12, .08]	.06
Deep Cohesion	.35**	.23**	.37*	.29	.08 [-.02, .26]	.22
Referential Cohesion	.35**	-.33**	.26*	.37**	-.12 [-.31, -.01]	.45

\*  $p < .10$ . \*\*  $p < .05$ .

Table 8  
*Partial Correlations Between Narrativity Subcomponents and Word Concreteness*

Measure	Partial correlations
Syllables per Word –	.45 <sup>a</sup>
Noun Incidence –	.68 <sup>a</sup>
Verb Incidence +	–.10 <sup>a</sup>
Adjective Incidence –	.35 <sup>a</sup>
Adverb Incidence +	–.54 <sup>a</sup>
Pronoun Incidence +	–.39 <sup>a</sup>
First-Person Pronoun Incidence +	–.53 <sup>a</sup>
Third-Person Pronoun Incidence +	.46
Word Frequency (logarithm) +	–.24 <sup>a</sup>
Content Word Frequency +	–.49 <sup>a</sup>
Age of Acquisition –	–.38
Minimum Word Frequency per Sentence +	–.46 <sup>a</sup>
Word Familiarity +	–.15 <sup>a</sup>
Negation Density +	–.13 <sup>a</sup>
Modifiers per Noun Phrase, mean –	.45 <sup>a</sup>
Passive Constructions Density –	.46 <sup>a</sup>

*Note.* Partial correlation = controlled for academic discipline. A plus (+) or minus (–) designate direction of loading on principal component. Because of the overall negative correlation between Word Concreteness and Narrativity, the subcomponents with positive loading are expected to correlate negatively with Word Concreteness, whereas the subcomponents with negative loading are expected to correlate positively with Word Concreteness.

<sup>a</sup> Consistent direction of correlation.

indicates an increase in word and/or sentence length—the features that can be assumed to increase text difficulty by having concepts of lower familiarity to the readers or by increasing working memory load (Caplan & Waters, 1999; Graesser et al., 2014; McNamara et al., 2014). Increased lexical diversity (e.g., MTLTD) and noun incidence can also add to text difficulty by increasing the density of information (Biber, 1988; McNamara et al., 2014). A frequent use of noun modifiers increases the readers’ working memory load, thus making a text more challenging (Graesser, Cai, Louwerse, & Daniel, 2006). Passive voice is acquired later in language development (Kirby, 2010) and can be more difficult to comprehend (Maratsos, Fox, Becker, & Chalkley, 1985) compared

Table 9  
*Correlations Between Deep Cohesion Subcomponents and Word Concreteness*

Measure	Partial correlations
Connectives Incidence +	–.16 <sup>a</sup>
Causal Connectives Incidence +	–.46 <sup>a</sup>
Temporal Connectives Incidence +	–.19 <sup>a</sup>
Logical Connectives Incidence +	–.52 <sup>a</sup>
Causal Cohesion +	–.31 <sup>a</sup>
Intentional Cohesion +	–.04 <sup>a</sup>

*Note.* Partial correlation = controlled for academic discipline. A plus (+) or minus (–) designate direction of loading on principal component. Because of the overall negative correlation between Word Concreteness and Deep Cohesion, the subcomponents with positive loading are expected to correlate negatively with Word Concreteness, whereas the subcomponents with negative loading are expected to correlate positively with Word Concreteness.

<sup>a</sup> Consistent direction of correlation.

Table 10  
*Correlations Between Syntactic Simplicity Subcomponents and Word Concreteness*

Measure	Partial correlations
Causal Verb Incidence +	–.26 <sup>a</sup>
Intentional Verb Incidence +	–.28 <sup>a</sup>
Words per Sentence –	.48 <sup>a</sup>
Words Before Main Verb of Main Clause –	.44 <sup>a</sup>
Syntactic Similarity +	–.34 <sup>a</sup>

*Note.* Partial correlation = controlled for academic discipline. A plus (+) or minus (–) designate direction of loading on principal component. Because of the overall negative correlation between Word Concreteness and Syntactic Simplicity, the subcomponents with positive loading are expected to correlate negatively with Word Concreteness, whereas the subcomponents with negative loading are expected to correlate positively with Word Concreteness.

<sup>a</sup> Consistent direction of correlation.

with active voice. Finally, TTR is another measure of lexical diversity. Higher TTR indicates lower cohesion, that is, a proportionally high use of unique words (types) in relation to the total number of words (tokens)—a feature that can increase text difficulty (Haberlandt & Graesser, 1985; McNamara & Kintsch, 1996).

**Regression analysis.** The average word concreteness/meaningfulness/imageability composite was higher in humanities than natural sciences. The same was true for FKGL, lexical diversity, modifiers per noun, noun incidence, passive voice incidence, and TTR. On the other hand, causal connectives incidence and content “word overlap – all sentences” were higher in natural sciences than humanities. Thus, academic discipline had a significant effect in the expected direction on all of the explored variables (i.e., natural sciences more difficult on concreteness/meaningfulness/imageability, but easier on the other measures than humanities).

**Mediation analysis.** In this set of mediation analyses, we used the same design as the previous analyses. Here, the individual

Table 11  
*Correlations Between Referential Cohesion Subcomponents and Word Concreteness*

Measure	Partial correlations
Content Word Overlap, Adjacent Sentences +	–.34 <sup>a</sup>
Content Word Overlap, All Sentences +	–.31 <sup>a</sup>
Argument Overlap, Adjacent Sentences +	.37
Argument Overlap, All Sentences +	.34
Noun Overlap, Adjacent Sentences +	.47
Stem Overlap, All Sentences +	.55
Type-Token Ratio –	.50 <sup>a</sup>
Lexical Diversity, All Words –	.58 <sup>a</sup>
Lexical Diversity, Verb Cohesion –	–.25
LSA, Given/New +	–.10 <sup>a</sup>
LSA Overlap, Adjacent Sentences +	.25
LSA Overlap, All Sentences +	.24

*Note.* Partial correlation = controlled for academic discipline. A plus (+) or minus (–) designate direction of loading on principal component. Because of the overall negative correlation between Word Concreteness and Referential Cohesion, the subcomponents with positive loading are expected to correlate negatively with Word Concreteness, whereas the subcomponents with negative loading are expected to correlate positively with Word Concreteness. LSA = latent semantic analysis.

<sup>a</sup> Consistent direction of correlation.

Table 12  
*Mean Values of Descriptive Measures, Principal Components, and Composite Measures Across Intervals, Yale and MIT Lectures Only*

Measure	Interval							
	1	2	3	4	5	6	7	8
Number of Words	677	691	703	700	693	682	705	688
Number of Sentences	36.7	37.8	38.6	39.7	37.6	39.7	40.0	40.4
Words per Sentence	20.3	19.9	19.6	19.0	19.7	18.6	18.7	18.2
Word Concreteness	-1.27	-1.35	-1.33	-1.41	-1.38	-1.35	-1.42	-1.48
Narrativity	.55	.61	.69	.62	.59	.64	.54	.56
Deep Cohesion	1.00	.77	.99	.88	.94	1.17	1.06	.77
Referential Cohesion	.05	.02	.08	-.01	-.02	.02	-.07	-.04
Syntactic Simplicity	-.27	-.14	-.13	-.05	-.15	-.05	-.03	.03
Overall Ease	.02	-.005	.09	.02	-.01	.07	-.01	.03
Formality	.42	.35	.39	.36	.37	.37	.35	.32

Note. MIT = Massachusetts Institute of Technology.

mediator was the average word concreteness/meaningfulness/imageability composite. The relation between academic discipline and causal connectives incidence, FKGL, lexical diversity, modifiers per noun, noun incidence, passive voice incidence, and TTR was mediated by the composite measure. The same was not true for the relation between academic discipline and content “word overlap – all sentences.” Thus, overall, evidence for linguistic compensation (i.e., as discourse difficulty indexed by word concreteness/meaningfulness/imageability increases, discourse difficulty on the majority of the other explored dimensions decreases) is apparent independent of the Coh-Matrix principal components.

**Within-lecture analysis.** Our last set of analyses explored how the average word concreteness/meaningfulness/imageability composite and FKGL change as a function of time in the lecture. We calculated slopes relating time to these measures and compared that value with zero (Yale and MIT lectures only). Whereas the average word concreteness/meaningfulness/imageability ( $-.17$  [ $-.60, .26$ ]) did not change across time,  $t(39) = -.81$ ,  $SEM = .21$ ,  $p = .42$ ,  $d = .13$ , FKGL ( $-.11$  [ $-.18, -.05$ ]) decreased across lecture intervals,  $t(39) = -3.40$ ,  $SEM = .03$ ,  $p = .002$ ,  $d = .52$ . When each of the three Word Concreteness subcomponents were analyzed separately, word meaningfulness decreased across lecture intervals ( $-.42$  [ $-.73, -.11$ ]),  $t(39) = -2.73$ ,  $SEM = .15$ ,  $p =$

$.009$ ,  $d = .44$ , whereas word concreteness ( $-.03$  [ $-.61, .55$ ]),  $t(39) = -.09$ ,  $SEM = .19$ ,  $p = .92$ ,  $d = .02$ , and word imageability, ( $-.07$  [ $-.60, .46$ ]),  $t(39) = -.07$ ,  $SEM = .26$ ,  $p = .26$ ,  $d = .04$ , did not change.

## General Discussion

The current study has provided a comprehensive test of the existence of linguistic compensation in the context of instructional oral discourse, specifically, lectures. The results revealed that the natural science lectures were more abstract (i.e., lower in word concreteness, meaningfulness, and imageability), but more narrative-like, cohesive, and syntactically simpler, compared with the humanities lectures. Although the results with respect to Narrativity were different than that found in texts, these results are consistent with the idea that decreases in content ease in the science lectures (as indexed by Word Concreteness) are compensated for by the use of more narrative-like, more cohesive, and syntactically simpler discourse. Previous work investigating linguistic compensation has typically focused on these more qualitative patterns across genres and academic disciplines to motivate the idea that authors modulate different dimensions of difficulty to

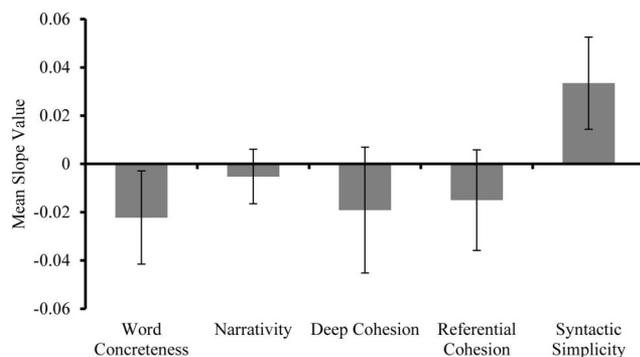


Figure 2. The principal components' mean slopes relating time into lecture and the five Coh-Matrix principal components. Error bars indicate the 95% confidence intervals.

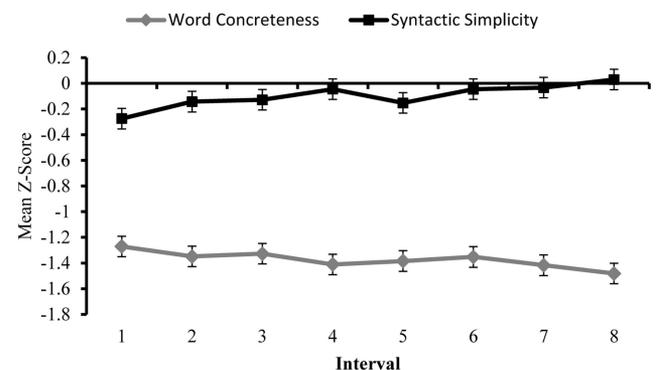


Figure 3. Mean values of Word Concreteness and Syntactic Simplicity across lecture intervals. Error bars indicate the 95% Loftus and Masson (1994) confidence intervals.

Table 13

*Regression Analyses, Individual Subcomponents as the Indicators of Linguistic Compensation, With Academic Discipline (0 = Humanities vs. 1 = Natural Sciences) as the Predictor, and the Subcomponents as the Outcome, Controlled for Academic Institution (Dummy Coded)*

Measure	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Word Concreteness composite	-14.43	3.88	-3.73	.000
Causal connectives incidence	4.40	1.87	2.36	.022
Content word overlap – all sentences	.01	.005	2.84	.007
FKGL	-1.58	.60	-2.16	.012
MTLD	-6.28	2.68	-2.34	.023
Modifiers per noun	-.06	.02	-2.32	.025
Noun incidence	-23.32	6.63	-3.52	.001
Passive voice incidence	-.96	.40	-2.78	.008
TTR	-.07	.01	-4.74	.000

*Note.* FKGL = Flesch-Kincaid Grade Level; MTLD = Measure of Textual Lexical Diversity; TTR = Type-Token Ratio.

accommodate their audience. In the present investigation, we took additional steps to test a compensation account.

Across a series of mediation analyses, we demonstrated that the relation between academic discipline and Narrativity, Deep Cohesion, and Syntactic Simplicity, but not Referential Cohesion, was mediated by Word Concreteness. In other words, the differences in these principal components across academic discipline could be explained (at least partially) by variation across lectures in Word Concreteness (i.e., word concreteness, meaningfulness, and imageability). This is consistent with the idea that lecturers were potentially compensating for the use of abstract words by simplifying other linguistic dimensions of oral discourse (i.e., more narrative-like, more cohesive, and syntactically simpler). The mediation analyses also demonstrated that, independent of academic discipline, Narrativity, Deep Cohesion, and Syntactic Simplicity were negatively correlated with Word Concreteness. Thus, there was evidence that within a discipline, lecturing about more abstract unfamiliar content was associated with using more narrative-like discourse, providing more connectives, and using simpler syntax. Although the primary motivation for the suggested causal ordering

is theoretical we also presented a series of alternative mediation models that provided further support for the idea of Word Concreteness driven compensation. However, these analyses of course do not rule out other potential compensatory relations between discourse dimensions, and cannot confirm our proposed causal ordering.

Although the mediation analyses provided evidence consistent with linguistic compensation (driven by Word Concreteness) with respect to Narrativity, Deep Cohesion, and Syntactic Simplicity, this was not the case for Referential Cohesion. Specifically, although Referential Cohesion was greater in the natural sciences than the humanities (marginally significant), the mediation analyses suggested that this difference was not attributable to differences across discipline in Word Concreteness. In a similar vein, when academic discipline was controlled, there was no relation between Word Concreteness and Referential Cohesion. The lack of evidence for Word Concreteness driven linguistic compensation with respect to Referential Cohesion is interesting, provided prior claims based on research using texts. We discuss this issue in the Relation to Previous Research section.

Importantly, we also provided a crucial demonstration that over time, within a lecture, Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) decreases, and that this is accompanied by increases in Syntactic Simplicity. Specifically, these analyses support the idea that compensation is an explicit or implicit cognitive process that is executed by an individual lecturer (as opposed to merely a fixed difference between lecturers or academic disciplines). In addition, the results of our analyses demonstrate that two composite measures can be used to explore overall ease or difficulty. In the following, we further discuss the basis for arguing that increasing content difficulty is driving linguistic compensation, the relation of the present work to previous research, the implications for Coh-Metrix as a tool to study linguistic processes, and the potential mechanism underlying the proposed linguistic compensation in lectures.

### Word Concreteness Driven Linguistic Compensation

The characterization of linguistic compensation that has emerged within the present investigation, similar to previous re-

Table 14

*Mediation Analyses: Average Word Concreteness/Meaningfulness/Imageability as the Mediator of the Relation Between Academic Discipline and the Individual Subcomponents, Controlled for Academic Institution (Dummy Coded)*

Outcome	Unstandardized regression coefficients				Indirect effect [95% CI]	Effect size
	a	b	c	c'		
Causal connectives	-14.4***	-.19**	4.40**	1.57	2.83 [.88, 6.70]	.64
Cont. word overlap	-14.4***	-.0004**	.01**	-.08	.005 [-.0001, .01]	.38
FKGL	-14.4***	.04**	-1.58**	-.93	-.65 [-1.97, -.04]	.41
MTLD	-14.4***	.29**	-6.28**	-2.13	-4.15 [-10.7, -.72]	.66
Modifiers per noun	-14.4***	.003	-.06**	-.007	-.05 [-.10, -.01]	.87
Noun incidence	-14.4***	1.16***	-23.3***	-6.49	-16.8 [-33, -6.3]	.72
Passive voice	-14.4***	.03*	-.96***	-.47	-.49 [-1.28, -.15]	.51
TTR	-14.4***	.001**	-.07***	-.05**	-.02 [-.05, -.005]	.27

*Note.* FKGL = Flesch-Kincaid Grade Level; MTLD = Measure of Textual Lexical Diversity; TTR = Type-Token Ratio.

\*  $p < .10$  level. \*\*  $p < .05$ . \*\*\*  $p < .001$ .

search (e.g., Graesser et al., 2011, 2014; McNamara, 2013), places special emphasis on the Word Concreteness principal component (consisting of three measures: word concreteness, word imageability, and word meaningfulness) as the driver of the compensating behavior (e.g., the increasing abstractness in the natural sciences leads lecturers to use simpler syntax). This claim is based on three important considerations. First, the linguistic variables that compose the Word Concreteness principal component have been demonstrated to be intimately tied to processing difficulty at a number of different levels. For example, concrete words are more quickly recognized, read, and recalled than abstract words (Allen & Hulme, 2006; Kroll & Merves, 1986; D. L. Nelson & Schreiber, 1992; Strain et al., 1995). Moreover, previous studies have demonstrated that participants recalled almost 3 times more content from concrete paragraphs than from abstract paragraphs (Paivio, 1990; Sadoski, Goetz, & Fritz, 1993). Sadoski et al. (1993) found a similar benefit for concrete paragraphs, despite controlling for length and readability across the texts (calculated by the FKGL formula). In addition, their study demonstrated that participants rated concrete sentences as more comprehensible and more interesting than abstract sentences. This latter finding is important in the context of other studies that have demonstrated that as text-based interest increases, recall and comprehension of the text increase (Hidi & Baird, 1988; Schraw et al., 1995; Schraw, Flowerday, & Lehman, 2001; Wade & Adams, 1990). Finally, Mosenthal (1996) demonstrated that participants rated questions that are higher in concreteness (e.g., asking information about persons, animals, or things) as easier to answer, compared with the questions asking more abstract information (e.g., information about comparison, contrast, reason, or cause). In a similar vein, the other two linguistic variables that contribute to the Word Concreteness principal component—word meaningfulness and word imageability—have similar effects on processing difficulty. For example, highly imageable words are named and recalled faster than words lower on this dimension (Bourassa & Besner, 1994; Strain et al., 1995), and meaningfulness enhances text comprehension (Danks, 1969). Thus, previous research has demonstrated that the linguistic variables that make up the Word Concreteness principal component are intimately tied to text interest, comprehensibility, and memorability.

Furthermore, previous studies (Reynolds & Paivio, 1968; Sadoski, Goetz, Stricker, & Burdinski, 2003; Sadoski, Kealy, Goetz, & Paivio, 1997; Sadoski & Paivio, 2007) have demonstrated that individuals vary their written and oral discourse composition features as a function of having to define more concrete or more abstract words. For example, when asked to define concrete concepts, individuals produced more words, more modifiers in sentences, and more fluid discourse (i.e., less hesitations and speech disfluencies) than when asked to define abstract concepts (Reynolds & Paivio, 1968; Sadoski & Goetz, 1998). According to DCT, concrete language induces more meaning and recall than the more abstract language because concrete language can be encoded as both language and imagery, whereas the more abstract language is mostly encoded as language but not imagery (Sadoski & Goetz, 1998; Sadoski et al., 2003).

Although the linguistic variables that make up the Word Concreteness principal component are intimately tied to difficulty broadly construed, in the case of informational discourse (e.g., lectures), the nature of the material often dictates the use of words

that would yield a low score on this dimension (McNamara, 2013). This leads to our final consideration motivating the proposal that abstractness can drive compensating behavior. Specifically, content domain (arguably) places more direct constraints on the discourse measures that make up the Word Concreteness principal component than it does on other aspects of language use in a lecture, leaving the latter to more freely vary. For example, lecturing about quantum physics likely places more constraints on the lecturer's words than it does on their syntax. As a result, at least in the context of informational discourse, it seems more reasonable to construct a theoretical case for differences across domain in Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) leading to differences across domain in Syntactic Simplicity, than to argue that differences across domain in Syntactic Simplicity lead to differences in Word Concreteness (i.e., word concreteness, meaningfulness, and imageability). This content domain constraint is of course tied directly to the fact that the linguistic variables that make up the Word Concreteness principal component are all word-level variables. That said, it is important to note that in analyzing natural speech in lectures, we have used a correlational approach, and, as such, the associated caveats of this approach (e.g., with regard to causation) need to be kept in mind.

The test of the alternative mediation models suggests that only the reverse causal model with Syntactic Simplicity is plausible when the other principal components were controlled. However, in the reverse causal model, Syntactic Simplicity did not mediate the relation between academic discipline and either Narrativity or Deep Cohesion, whereas, of course, Word Concreteness did. Thus, again, with the proper caveats in mind regarding correlational data, the analyses seem to support the Word Concreteness driven linguistic compensation hypothesis. The focus here on Word Concreteness driven linguistic compensation does not preclude other potential compensatory relations between different dimensions of text/speech difficulty. Future research exploring such relations will represent an important contribution to further understanding potential varieties of linguistic compensation.

One of the primary reasons that we have put forward the Word Concreteness principal component as the driver of compensation is that the three word-level indices that make it up are strongly tied to the content and should, as a result, be more constrained by topic. The Narrativity principal component, to some extent, also has this feature. For example, it contains measures of word frequency and familiarity, and higher Narrativity usually indicates discourse with familiar topics (Biber, 1988; Graesser et al., 2011). Thus, it was surprising, given previous research (Graesser et al., 2011; McNamara, 2013), that, in this corpus, the two principal components were negatively correlated. Given the strong ties between Word Concreteness, Narrativity, and content, an interesting alternative approach in investigating linguistic compensation could involve, a priori, defining "content difficulty" as a composite of Word Concreteness and Narrativity. The investigation would examine the extent to which writers/lecturers compensate for changes in this index of content difficulty using the remaining three dimensions (i.e., Deep Cohesion, Referential Cohesion, and Syntactic Simplicity). These latter measures are putatively freer to vary across levels of content difficulty (and, as such, represent ideal targets for compensating behavior). Using this measure, discourse about unfamiliar topics that contains more abstract words would be con-

sidered more difficult than discourse about familiar topics with more concrete words on the other end of this measure. Research investigating this composite is currently underway.

### Gradients of Evidence for Compensation?

We have considered three sources of empirical evidence for the notion that individuals are compensating for changes in Word Concreteness by altering other linguistic features of the lecture (i.e., Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity): (a) a contrasting pattern of means across academic discipline (e.g., in which Word Concreteness is lower the other dimension is higher), (b) the effect of academic discipline on a given dimension is mediated by Word Concreteness, and (c) contrasting patterns of change within lecture (i.e., given Word Concreteness decreases across the lecture, compensation would be evidenced by an increase in scores on a given dimension). Although meeting the first test and not the others would clearly suggest a noncompensatory argument to explain (a), it is interesting to consider the implications of meeting (b) but not (c), as in the case of Narrativity and Deep Cohesion, versus meeting both (b) and (c), as in the case of Syntactic Simplicity. If we use the presence of a within-lecture pattern of compensation in addition to the evidence for mediation as the evidentiary lower limit for the presence of linguistic compensation, then only Syntactic Simplicity crosses that bar. There are good a priori grounds for expecting that it would be effective to compensate for content difficulty by simplifying syntax. For example, longer sentences place more burden on working memory and are harder to process compared with shorter sentences (e.g., McNamara et al., 2014). Indeed, traditional metrics of readability, such as FKGL and Flesch-Kincaid Reading Ease (Klare, 1974–1975), are largely measures of sentence length (and word length). Lastly, increasing Syntactic Simplicity as a means to increase comprehensibility is likely a technique most academics (i.e., the instructors in our lectures and book publishers) will have encountered (e.g., in their own writing). Thus, in the context of linguistic compensation, simplifying syntax represents a sensible choice to offset discourse difficulty on another dimension (e.g., lower word concreteness, meaningfulness, and imageability).

Interpreting the evidence presented here as support for Word Concreteness driven compensation of Syntactic Simplicity makes for a simple and straightforward characterization of linguistic compensation. However, there is good reason to think this latter position might be too conservative. One point to this effect is that the 40-min lecture transcripts in our corpus offer a relatively narrow range of scores within lecture. This makes the detection of within-lecture variations in principal components difficult. For example, within-lecture Word Concreteness varied from a  $z$  value of 1.27 in the first 5 min to a value of 1.48 in the last 5 min. On the other hand, across academic discipline, Word Concreteness varied from a  $z$  value of  $-.88$  in the humanities to a value of  $-1.93$  in the natural sciences. Thus, the difference in Word Concreteness across academic discipline is about 5 times that found within a lecture, arguably offering a greater opportunity for linguistic compensation to emerge. In addition, there are likely other variables constraining the amount and direction of change in the use of language throughout a lecture. Given the dearth of research evaluating such change, this seems an excellent opportunity for

future research. Lastly, interpreting the evidence presented here as support for Word Concreteness driven compensation of Syntactic Simplicity only would leave the evidence from mediation unexplained in the case of Narrativity and Deep Cohesion. However, these latter relations could be related to some currently unknown third variable that is driving changes on both dimensions. A potentially more productive way (at least at present) to view the pattern across the different measures is to assume that there is evidence of Word Concreteness driven linguistic compensation in Narrativity, Deep Cohesion, and Syntactic Simplicity, with the evidence being stronger in the latter case than the former two.

### Relation to Previous Research

Although the variations in discourse features across academic discipline reported here are largely consistent with previous studies (e.g., Graesser et al., 2011, 2014; McNamara, 2013), there were potentially important differences as well. For example, McNamara (2013) demonstrated that writers of natural science texts used lower Word Concreteness and Narrativity, but simpler syntax and higher Referential Cohesion, compared with the writers in language arts. Our analyses of lectures revealed a somewhat different pattern. Namely, whereas natural science lectures (like texts) were lower in Word Concreteness and higher in Syntactic Simplicity and Referential Cohesion (marginally), they were (unlike texts) more narrative-like (rather than less) and higher in Deep Cohesion, compared with the humanities lectures. There are a number of differences between the previous work and the current study that could explain variations across studies. The most salient difference, of course, is that previous studies (Graesser et al., 2011, 2014; McNamara, 2013) focused on texts within the TASA corpus, which, in the language arts classification, included actual narratives. Thus, in a comparison between language arts and natural sciences, we would expect Narrativity to be higher in the former (as was observed in that work; Graesser et al., 2011; McNamara, 2013). In the humanities lectures used here, there were no actual narratives (the only exception being proportionally small parts of lectures in which lecturers read literal excerpts, poems, etc.). In other words, all of the lectures were “informational,” talking about narratives rather than the telling of stories or other forms of narrative. Thus, in the present sample, the intent of the authors (i.e., lecturers) is more homogenous and academic than in previous work. Critically, despite this homogeneity, we still observed evidence for compensation and a pattern largely consistent with previous research in texts. In addition, our results indicated that natural science lectures were higher in formality compared with the humanities lectures. Previous research (Graesser et al., 2014; Li et al., 2013) has reported similar results when texts in different genres were analyzed. For example, language arts texts were less formal compared with natural science texts. Our results support the idea that formality can be used as a singular Coh-Matrix dimension of text difficulty (Graesser et al., 2014). With respect to a singular dimension of difficulty, our measure of overall ease (i.e., the average of all component scores) also yielded interesting patterns, both across academic discipline and within lecture. In particular, the natural sciences were higher in overall ease than the humanities, despite the use of more abstract, less familiar, and less meaningful words, and there was no change in overall ease across

time in the lecture, again despite the more frequent use of abstract, less familiar, and less meaningful words as time into the lecture increased. As noted in the introduction, one way of conceptualizing linguistic compensation within an educational context is to suggest that instructors are using it to stay within an optimal zone of comprehension difficulty (i.e., not too easy, not too hard, e.g., Graesser et al., 2014; McNamara et al., 2014; Wolfe et al., 1998). In this sense, if individuals are compensating for the difficult content (as measured by the Word Concreteness principal component), then this effort in a linguistic (and Coh-Matrix) sense was largely successful (i.e., the less concrete natural sciences were “easier” overall and there was no change, despite the decreasing concreteness, in ease across the lecture). Future work exploring these composite measures will provide a better understanding of the extent to which they index meaningful aspects of discourse.

The present study has also directly tested the implied causal relations between variations in Coh-Matrix dimensions reported in previous investigations (Graesser et al., 2011; McNamara, 2013). Our results clearly supported the need for such analysis when testing the compensation account of linguistic patterns across genre/academic disciplines. For example, in the present study, the natural science lectures had more difficult content and were more cohesive compared with the humanities lectures—a pattern of means consistent with the compensation account (e.g., the use of more abstract words is compensated for by using more cohesive language). However, we found no evidence (beyond the marginally significant pattern of means) that would support such an interpretation for the Referential Cohesion component. For example, the difference between academic disciplines in Referential Cohesion was not because of the differences between academic disciplines in Word Concreteness. Furthermore, after controlling for academic discipline, there was no correlation between Referential Cohesion and Word Concreteness. This failure to find evidence of compensation in terms of Referential Cohesion is interesting theoretically, given the claim that such compensation exists in texts (Graesser et al., 2011; McNamara, 2013).

There exist a number of potential explanations for this putative difference in results for Referential Cohesion. First, there may be no difference at all in the sense that if the mediation analyses performed here could have been conducted in previous studies, evidence for compensation with respect to Referential Cohesion would not have been found in text. Second, it is possible that the Referential Cohesion principal component is easier to manipulate in written than in oral discourse, or that the Referential Cohesion subcomponents behave differently across spoken and written registers (i.e., there is a difference across medium). Print differs from oral discourse by allowing the author to edit text rather than constructing the text extemporaneously. Interestingly, when student writing is time-pressured and the possibility of editing limited, cohesion has a diminished role in analysis (McNamara et al., 2010). Third, the difference reflects, in some manner, the use of actual narratives in the analysis of texts versus the focus here on informational discourse. Future research that can more directly assess differences between written and oral instruction is needed to address this question.

## Implications for the Coh-Matrix Principal Components

The present research focused on five of the Coh-Matrix principal components (Word Concreteness, Narrativity, Deep Cohesion, Referential Cohesion, and Syntactic Simplicity). Although our goal was not to evaluate these components, the subcomponent analysis revealed both support for the components and a potential challenge, specifically for the Referential Cohesion principal component. The results of our analysis revealed that, in general, principal components’ subcomponents correlated with Word Concreteness in the direction expected by the Word Concreteness driven compensatory account (which is also the direction of the overall zero-order correlation between Word Concreteness and the principal components). This was generally true for the Narrativity, Syntactic Simplicity, and the Deep Cohesion subcomponents. However, in the case of the Referential Cohesion subcomponents, about half did not correlate with Word Concreteness in the expected direction (i.e., argument, stem, noun and LSA overlap measures), and about half did (i.e., content word overlap, TTR, LSA given/new and lexical diversity). Importantly, as is evident in Table 11, it does not appear to be the case that some components correlate in the expected direction and others are simply uncorrelated; rather, it appears that some components are correlated in one direction and others in the opposite direction. This of course does not imply that the Referential Cohesion principal component should not be considered when exploring linguistic compensation. One possible explanation is that the subcomponents that make up the Referential Cohesion principal component behave differently in written versus oral discourse (the principal components were generated using written text; Graesser et al., 2011). Clearly, more work is needed to better understand the apparent heterogeneity of the Referential Cohesion principal component in oral discourse. In addition, it seems clear that when investigating the Coh-Matrix components, it is important to also assess the behavior of the subcomponents.

Even though some of our results indicated potential difficulties with one of the principal components in Coh-Matrix, overall, the present investigation further supports the utility of such an approach. Specifically, the other principal components behaved in a systematic fashion (and as analysis of the subcomponents indicate, in a rather homogenous manner) and afforded a number of important insights into potential linguistic sources of lecture difficulty. It is also important to note that although the present investigation has relied heavily on the Coh-Matrix principal components, the notion of linguistic compensation is certainly not tied to this particular grouping of variables. For example, a researcher could instead select individual variables (e.g., word concreteness, sentence length, incidence of causal connectives, lexical diversity, verb incidence) and conduct similar analyses without the theoretical commitments made within the Coh-Matrix framework. The subcomponents analysis suggests that the patterns would be largely the same, although it could reveal further nuances obscured by collapsing across multiple linguistic variables. For example, such an approach might be helpful in better understanding of linguistic compensation with respect to the Referential Cohesion subcomponent (i.e., some of the subcomponents do correlate in the expected direction with Word Concreteness) and negative correlation between Word Concreteness and Narrativity that would be expected to positively correlate.

## Variations in Time in Lectures

Previous research (Lightman et al., 2007) suggested that text difficulty, indexed by FKGL, increased across chapters in history and science textbooks, whereas cohesion remained unchanged. In a similar vein, our analyses of the changes in principal components across time in lectures revealed that lecture ease (as indexed by Word Concreteness) decreased as a function of time. Interestingly, and consistent with the Word Concreteness driven compensation account, this decrease in Word Concreteness (i.e., word concreteness, meaningfulness, and imageability) across time in the lecture was accompanied by an increase in Syntactic Simplicity. To our knowledge, this is the first demonstration of such linguistic variations within a lecture. Interestingly, on the other side of the lectern, researchers also find variations over time. In particular, as time increases, mind wandering increases and memory for the material presented decreases (Farley, Risko, & Kingstone, 2013; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Risko, Buchanan, Medimorec, & Kingstone, 2013). Although this pattern has typically been attributed to a failing in the attentional system of the student, it is interesting to consider the potential contribution of variations in the difficulty of the discourse to student's attention. For example, one explanation for the increase in mind wandering over time in a lecture is that students' situation model of the lecture breaks down and, as a result, students disengage (Risko et al., 2013; Smallwood, 2011; Smallwood, Fishman, & Schooler, 2007). A decrease in Word Concreteness over time (that is not compensated for) would arguably increase the likelihood of a situation model breaking down. In addition, it will be important to consider, in future work, how subtle cues from the audience (i.e., the students) might modify lecturer language (e.g., do they simplify their language when students look puzzled?). A deeper understanding of the interaction between student and lecturer could be provided through combining the types of analyses here with paradigms developed to explore mind wandering.

## Potential Mechanisms Underlying Linguistic Compensation

According to previous research (e.g., Graesser et al., 2011, 2014; McNamara, 2013), authors of texts purposefully engage in compensating linguistic behavior because of their epistemic stance toward the intended audience. The results of the present investigation suggest that lecturers, similar to authors of texts, offset content difficulty by varying different components of discourse. This compensation behavior can be considered a kind of metalinguistic act. According to Tunmer, Herriman, and Nesdale (1988), individuals rely on metalinguistic ability to "reflect on and manipulate the structural features of spoken language" (p. 136), and this ability relies on controlled processing. In the case of lecturers, this deliberate manipulation could come at the time of the preparation of lecture and/or as a response to students during a lecture. On this account, the compensation patterns found reflect a deliberate manipulation of discourse by authors of texts and lecturers. For example, provided that novel information (i.e., lecture content) requires more cognitive resources compared with the familiar information (Graesser & McNamara, 2011), lecturers may simplify discourse in order to accommodate students' (perceived) processing limitations. Arguably, engaging in this kind of inten-

tional compensatory linguistic behavior based on controlled processing might be much harder in the oral discourse context compared with writing, given the temporal demands. Nevertheless, the present investigation clearly suggests lecturers do so. This account of the linguistic patterns observed here makes a number of testable predictions. For example, if compensation is a controlled process, it should itself be subject to capacity limitations. In addition, if compensation represents a kind of metalinguistic act, then manipulations of the lecturer's metacognitive knowledge (e.g., prior knowledge of audience) should influence the compensatory behavior. In a similar vein, to the extent that linguistic compensation reflects a strategic act under the control of the instructor, we might also expect variation across language or culture to influence the extent to which lecturers compensate. While on this topic, it is worth stressing that our corpus consisted of lecture transcripts from three U.S. universities. Whether the compensatory patterns reported here represent idiosyncrasies of instructional English language and/or the U.S. educational system or they generalize broadly remains a question for future investigation.

A different account of the linguistic patterns observed here is that they are a byproduct of demands on production. In other words, the increased computational demands on selection, or planning discourse, likely associated with talking about difficult concepts (e.g., more abstract) might itself constrain the complexity of the utterance. Previous studies have demonstrated that increased cognitive load affects language production. For example, Müller et al. (2001) reported that speakers' utterances shortened with increases in cognitive load. Interestingly, the increased cognitive load did not have a large effect on content quality (as assigned by four independent raters). If the use of more abstract concepts in lectures increases the cognitive load on the lecturer, then this might, for example, force the speaker to adopt simpler syntax. This account would not require the lecturer to adopt an "epistemic stance" toward their audience or intentionally enact any efforts to "compensate" for difficulties. Rather, the observed compensatory linguistic behavior would "fall out" of the cognitive limitations of the lecturer (i.e., they are compensating for their own cognitive limitations rather than that of the audience).

In a similar vein, DCT (Clark & Paivio, 1991; Paivio, 1971, 1986; Sadoski & Paivio, 2013) might also be used to explain the apparent compensatory linguistic behavior. DCT argues for the existence of two separate, but interacting, cognitive systems in language processing and composition—a verbal system and a nonverbal (imagery) system. Thus, a verbal representation can potentially be associated with the other concepts within the verbal system, while also being linked with the corresponding nonverbal representations (i.e., images). In the context of the Word Concreteness driven linguistic compensation, concrete language (as opposed to abstract language) is more likely to evoke both the associated verbal representations and imagery. In this case, the images would offer another source of content (in addition to linguistic information). On the other hand, the more abstract content relies heavily on verbal strategies, because the number of referential (i.e., word to image) connections is limited (the lack of object referents). Indeed, in previous studies, participants consistently reported using mental imagery while defining concrete words, and a verbal strategy while defining the abstract words (e.g., Sadoski & Goetz, 1998). As a result, and as demonstrated by the previous research, concrete words generate more voluble def-

initions compared with abstract words (Reynolds & Paivio, 1968; Sadoski et al., 1997, 2003). For example, definitions of concrete words had shorter latencies (i.e., began sooner), were longer, and used more modifiers per sentence compared with definitions of abstract words. Similarly, our results indicate that the more concrete language in the humanities had higher noun incidence and more modifiers per noun phrase (i.e., lower Narrativity), more lexically diverse language (i.e., lower Referential Cohesion), and longer sentences (i.e., lower Syntactic Simplicity) than the more abstract language in the natural sciences. Thus, the apparent off-setting patterns reported here could reflect interactions between verbal and imagistic sources of information in oral composition.

Of course, these two potential mechanisms (controlled process vs. a natural byproduct of language production) are not mutually exclusive, and it is possible that a combination of both can be used to explain linguistic compensation in lectures.

### Conclusion

In conclusion, the current study represents one of the first broad tests of the existence of linguistic compensation in the context of oral discourse. Previous studies explored qualitative patterns of text principal components across genres/academic disciplines (e.g., Graesser et al., 2011, 2014; McNamara, 2013), and the present investigation has expanded this analysis to oral discourse in lectures and provided more direct tests of the compensation account. Results strongly suggest that linguistic compensation for sources of difficulty in discourse represents a common feature of language use in instructional contexts.

### References

- Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*, 64–88. <http://dx.doi.org/10.1016/j.jml.2006.02.002>
- Ashburn, G., & Gordon, A. (1981). Features of a simplified register in speech to elderly conversationalists. *International Journal of Psycholinguistics*, *8*, 7–31.
- Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, *26*, 251–276. <http://dx.doi.org/10.2307/747763>
- Beck, I. L., McKeown, M. G., & Worthy, J. (1995). Giving a text voice can improve students' understanding. *Reading Research Quarterly*, *30*, 220–238. <http://dx.doi.org/10.2307/748033>
- Biber, D. (1988). *Variation across speech and writing*. Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511621024>
- Bligh, D. A. (1972). *What's the use of lectures?* Baltimore, MD: Penguin Books.
- Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review*, *1*, 122–125. <http://dx.doi.org/10.3758/BF03200768>
- Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, *83*, 329–345. <http://dx.doi.org/10.1037/0022-0663.83.3.329>
- Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*, 429–441. <http://dx.doi.org/10.1037/a0022824>
- Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77–94. <http://dx.doi.org/10.1017/S0140525X99001788>
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210. <http://dx.doi.org/10.1007/BF01320076>
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *33*, 497–505. <http://dx.doi.org/10.1080/14640748108400805>
- Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication*, *10*, 39–71. <http://dx.doi.org/10.1177/0741088393010001002>
- Danks, J. H. (1969). Grammaticalness and meaningfulness in the comprehension of sentences. *Journal of Verbal Learning & Verbal Behavior*, *8*, 687–696. [http://dx.doi.org/10.1016/S0022-5371\(69\)80031-9](http://dx.doi.org/10.1016/S0022-5371(69)80031-9)
- Farley, J., Risko, E. F., & Kingstone, A. (2013). Everyday attention and lecture retention: The effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*, 619. <http://dx.doi.org/10.3389/fpsyg.2013.00619>
- Finkelstein, M. J., Seal, R. K., & Schuster, J. H. (1998). *The new academic generation: A profession in transformation*. Baltimore, MD: JHU Press.
- Fraser, C., & Roberts, N. (1975). Mothers' speech to children of four different ages. *Journal of Psycholinguistic Research*, *4*, 9–16. <http://dx.doi.org/10.1007/BF01066986>
- Glass, J., Hazen, T. J., Hetherington, L., & Wang, C. (2004). Analysis and processing of lecture audio data: Preliminary investigations. In: *Proceedings of the Workshop on Interdisciplinary Approaches to Speech Indexing and Retrieval at HLT-NAACL 2004* (pp. 9–12). Stroudsburg, PA: Association for Computational Linguistics.
- Graesser, A. C., Cai, Z., Louwerse, M. M., & Daniel, F. (2006). Question Understanding Aid (QUAID): A web facility that tests question comprehensibility. *Public Opinion Quarterly*, *70*, 3–22. <http://dx.doi.org/10.1093/poq/nfj012>
- Graesser, A. C., Hautt-Smith, K., Cohen, A. D., & Pyles, L. D. (1980). Advanced outlines, familiarity, text genre, and retention of prose. *Journal of Experimental Education*, *48*, 209–220.
- Graesser, A. C., Jeon, M., & Cai, Z. (2008, October 20). Presidential and vice-presidential candidates: Language beyond the word. *Wordwatchers*. Retrieved from <http://wordwatchers.wordpress.com/2008/10/20/presidential-and-vice-presidential-candidates-language-beyond-the-word/>
- Graesser, A. C., Jeon, M., Yan, Y., & Cai, Z. (2007). Discourse cohesion in text and tutorial dialogue. *Information Design Journal*, *15*, 199–213. <http://dx.doi.org/10.1075/idj.15.3.02gra>
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, *3*, 371–398. <http://dx.doi.org/10.1111/j.1756-8765.2010.01081.x>
- Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, *115*, 210–229. <http://dx.doi.org/10.1086/678293>
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix providing multilevel analyses of text characteristics. *Educational Researcher*, *40*, 223–234. <http://dx.doi.org/10.3102/0013189X11413260>
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments & Computers*, *36*, 193–202. <http://dx.doi.org/10.3758/BF03195564>
- Griffin, R. W., & Cashin, W. E. (1989). The lecture and discussion method for management education: Pros and cons. *Journal of Management Development*, *8*, 25–32. <http://dx.doi.org/10.1108/EUM000000001339>
- Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General*, *114*, 357–374. <http://dx.doi.org/10.1037/0096-3445.114.3.357>

- Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. Retrieved from <http://www.afhayes.com/public/process2012.pdf>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.
- Hidi, S., & Baird, W. (1988). Strategies for increasing text-based interest and students' recall of expository texts. *Reading Research Quarterly*, 23, 465–483. <http://dx.doi.org/10.2307/747644>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23–31. [http://dx.doi.org/10.1207/S15326985EP3801\\_4](http://dx.doi.org/10.1207/S15326985EP3801_4)
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.
- Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 14, 196–214. [http://dx.doi.org/10.1016/S0022-5371\(75\)80065-X](http://dx.doi.org/10.1016/S0022-5371(75)80065-X)
- Kirby, S. (2010). Passives in first language acquisition: What causes the delay? *University of Pennsylvania Working Papers in Linguistics*, 16, 107–117.
- Klare, G. R. (1974–1975). Assessing readability. *Reading Research Quarterly*, 10, 62–102. <http://dx.doi.org/10.2307/747086>
- Koedinger, K. R., & Alevan, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19, 239–264. <http://dx.doi.org/10.1007/s10648-007-9049-0>
- Koedinger, K. R., Pavlik, P., McLaren, B. M., & Alevan, V. (2008). Is it better to give than to receive? The assistance dilemma as a fundamental unsolved problem in the cognitive science of learning and instruction. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2155–2160). Austin, TX: Cognitive Science Society.
- Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 92–107. <http://dx.doi.org/10.1037/0278-7393.12.1.92>
- Lammers, W. J., & Murphy, J. J. (2002). A profile of teaching techniques used in the university classroom a descriptive profile of a us public university. *Active Learning in Higher Education*, 3, 54–67. <http://dx.doi.org/10.1177/1469787402003001005>
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284. <http://dx.doi.org/10.1080/01638539809545028>
- Li, H., Cai, Z., & Graesser, A. C. (2013, May). Comparing two measures for formality. In *FLAIRS Conference* (pp. 220–225). Palo Alto, California: AAAI Press.
- Lightman, E. J., McCarthy, P. M., Dufty, D. F., & McNamara, D. S. (2007). Cohesion and structural organization in high school texts. In D. Wilson & G. Sutcliffe (Eds.), *FLAIRS Conference* (pp. 235–240). Menlo Park, California: The AAAI Press.
- Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476–490. <http://dx.doi.org/10.3758/BF03210951>
- Louwerse, M. M., McCarthy, P. M., McNamara, D. S., & Graesser, A. C. (2004). Variation in language and cohesion across written and spoken registers. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 843–848). Mahwah, NJ: Erlbaum.
- MacDonald, M. C. (1997). Lexical representations and sentence processing: An introduction. *Language and Cognitive Processes*, 12, 121–136. <http://dx.doi.org/10.1080/016909697386826>
- MacKinnon, D. P. (1994). Analysis of mediating variables in prevention and intervention research. In A. Cazares & L. A. Beatty (Eds.), *Scientific methods for prevention intervention research* (pp. 127–153). Rockville, MD: National Institute on Drug Abuse.
- MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, 17, 144–158. <http://dx.doi.org/10.1177/0193841X9301700202>
- Magnifico, A. M. (2010). Writing for whom? Cognition, motivation, and a writer's audience. *Educational Psychologist*, 45, 167–184. <http://dx.doi.org/10.1080/00461520.2010.493470>
- Maratsos, M., Fox, D. E., Becker, J. A., & Chalkley, M. A. (1985). Semantic restrictions on children's passives. *Cognition*, 19, 167–191. [http://dx.doi.org/10.1016/0010-0277\(85\)90017-4](http://dx.doi.org/10.1016/0010-0277(85)90017-4)
- McCarthy, P. M., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42, 381–392. <http://dx.doi.org/10.3758/BRM.42.2.381>
- McNamara, D. S. (2013). The epistemic stance between the author and reader: A driving force in the cohesion of text and writing. *Discourse Studies*, 15, 579–595. <http://dx.doi.org/10.1177/1461445613501446>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, 27, 57–86. <http://dx.doi.org/10.1177/0741088309351547>
- McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2012). Sources of text difficulty: Across genres and grades. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 89–116). Plymouth, UK: Rowman & Littlefield Education.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511894664>
- McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247–288. <http://dx.doi.org/10.1080/01638539609544975>
- Mosenthal, P. (1996). Understanding the strategies of document literacy and their conditions of use. *Journal of Educational Psychology*, 88, 314–332. <http://dx.doi.org/10.1037/0022-0663.88.2.314>
- Müller, C., Großmann-Hutter, B., Jameson, A., Rummer, R., & Wittig, F. (2001). Recognizing time pressure and cognitive load on the basis of speech: An experimental study. *Lecture Notes in Computer Science*, 2109, 24–33. [http://dx.doi.org/10.1007/3-540-44566-8\\_3](http://dx.doi.org/10.1007/3-540-44566-8_3)
- Nelson, D. L., & Schreiber, T. A. (1992). Word Concreteness and word structure as independent determinants of recall. *Journal of Memory and Language*, 31, 237–260. [http://dx.doi.org/10.1016/0749-596X\(92\)90013-N](http://dx.doi.org/10.1016/0749-596X(92)90013-N)
- Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*. New York, NY: Student Achievement Partners.
- O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, 43, 121–152. <http://dx.doi.org/10.1080/01638530709336895>
- Paivio, A. (1971). *Imagery and verbal processes*. New York, NY: Holt, Rinehart & Winston.
- Paivio, A. (1990). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195066661.001.0001>
- Pearson, P. D., & Hiebert, E. F. (2010). National reports in literacy: Building a scientific base for practice and policy. *Educational Researcher*, 39, 286–294. <http://dx.doi.org/10.3102/0013189X10370205>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–891. <http://dx.doi.org/10.3758/BRM.40.3.879>
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, 16, 93–115. <http://dx.doi.org/10.1037/a0022658>

- Pritchard, D. (2010). Where learning starts? A framework for thinking about lectures in university mathematics. *International Journal of Mathematical Education in Science and Technology*, *41*, 609–623. <http://dx.doi.org/10.1080/00207391003605254>
- Reynolds, A., & Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Psychology*, *22*, 164–175.
- Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology*, *26*, 234–242. <http://dx.doi.org/10.1002/acp.1814>
- Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers & Education*, *68*, 275–283. <http://dx.doi.org/10.1016/j.compedu.2013.05.001>
- Sadoski, M., & Goetz, E. T. (1998). Concreteness effects and syntactic modification in written composition. *Scientific Studies of Reading*, *2*, 341–352. [http://dx.doi.org/10.1207/s1532799xssr0204\\_3](http://dx.doi.org/10.1207/s1532799xssr0204_3)
- Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). Impact of concreteness on comprehensibility, interest, and memory for text: Implications for dual coding theory and text design. *Journal of Educational Psychology*, *85*, 291–304. <http://dx.doi.org/10.1037/0022-0663.85.2.291>
- Sadoski, M., Goetz, E. T., Stricker, A. G., & Burdinski, T. K., Jr. (2003). New findings for concreteness and imagery effects in written composition. *Reading and Writing*, *16*, 443–453. <http://dx.doi.org/10.1023/A:1024235831671>
- Sadoski, M., Kealy, W. A., Goetz, E. T., & Paivio, A. (1997). Concreteness and imagery effects in the written composition of definitions. *Journal of Educational Psychology*, *89*, 518–526. <http://dx.doi.org/10.1037/0022-0663.89.3.518>
- Sadoski, M., & Paivio, A. (2007). Toward a unified theory of reading. *Scientific Studies of Reading*, *11*, 337–356. <http://dx.doi.org/10.1080/10888430701530714>
- Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing*. New York, NY: Routledge.
- Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology*, *60*, 362–367. <http://dx.doi.org/10.1037/0022-3514.60.3.362>
- Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 123–164). Hillsdale, NJ: Erlbaum.
- Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. *Journal of Reading Behavior*, *27*, 1–17. <http://dx.doi.org/10.1080/10862969509547866>
- Schraw, G., Flowerday, T., & Lehman, S. (2001). Increasing situational interest in the classroom. *Educational Psychology Review*, *13*, 211–224. <http://dx.doi.org/10.1023/A:1016619705184>
- Smallwood, J. (2011). The footprints of a wandering mind: Further examination of the time course of an attentional lapse. *Cognitive Neuroscience*, *2*, 91–97. <http://dx.doi.org/10.1080/17588928.2010.537746>
- Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*, *14*, 230–236. <http://dx.doi.org/10.3758/BF03194057>
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312. <http://dx.doi.org/10.2307/270723>
- Stein, D. (1992). *Cooperating with written text: The pragmatics and comprehension of written texts*. Berlin, Germany: De Gruyter.
- Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154. <http://dx.doi.org/10.1037/0278-7393.21.5.1140>
- Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, *58*, 1–21. <http://dx.doi.org/10.2307/413530>
- Templin, M. (1957). *Certain language skills in children: Their development and interrelationships*. Minneapolis, MN: University of Minnesota Press.
- Thornton, R., & Light, L. L. (2006). Language comprehension and production in normal aging. *Handbook of the psychology of aging*, *6*, 261–287. <http://dx.doi.org/10.1016/B9-78-012101-2/64950-015X>
- Tunmer, W. E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly*, *23*, 134–158. <http://dx.doi.org/10.2307/747799>
- Vidal-Abarca, E., Martínez, G., & Gilabert, R. (2000). Two procedures to improve instructional text: Effects on memory and learning. *Journal of Educational Psychology*, *92*, 107–116. <http://dx.doi.org/10.1037/0022-0663.92.1.107>
- Wade, S. E., & Adams, R. B. (1990). Effects of importance and interest on recall of biographical text. *Journal of Reading Behavior*, *22*, 331–353. <http://dx.doi.org/10.1080/10862969009547717>
- Walker, E. J., Risko, E. F., & Kingstone, A. (2014). Fillers as signals: Evidence from a question-answering paradigm. *Discourse Processes*, *51*, 264–286. <http://dx.doi.org/10.1080/0163853X.2013.862478>
- Watts, M., & Schaur, G. (2011). Teaching and assessment methods in undergraduate economics: A fourth national quinquennial survey. *The Journal of Economic Education*, *42*, 294–309. <http://dx.doi.org/10.1080/00220485.2011.581956>
- Wolfe, M. B., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). Learning from text: Matching readers and texts by latent semantic analysis. *Discourse Processes*, *25*, 309–336. <http://dx.doi.org/10.1080/01638539809545030>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185. <http://dx.doi.org/10.1037/0033-2909.123.2.162>

Received May 14, 2014

Revision received December 1, 2014

Accepted December 4, 2014 ■