
A Text Relatedness and Dependency Computational Model:

Using Latent Semantic Analysis and Coh-Metrix to
Predict Self-explanation Quality
Running
head: Relatedness and Text Dependency Model

Cédric Bellissens^{*} — Patrick Jeuniaux^{**} — Nicholas D. Duran^{***}
— Danielle S. McNamara^{****}

^{*} Institute for Intelligent Systems
University of Memphis, Memphis, TN USA

^{*} cbellissens@mail.psyc.memphis.edu

^{**} pjeuniaux@mail.psyc.memphis.edu

^{***} nduran@mail.psyc.memphis.edu

^{****} d.mcnamara@mail.psyc.memphis.edu

ABSTRACT. *The Interactive Strategy Trainer for Active Reading and Thinking (iSTART) is an intelligent tutoring system that provides students with automated training on reading strategies. In particular, iSTART helps students integrate textual information into a coherent mental representation through self-explanation. The goal of the present study was to examine how text cohesion influences qualitatively different types of self-explanation, namely, bridging and elaborative inferences. To do so, we developed a computational model that characterizes cohesion in terms of the textbase indices word stem and Latent Semantic Analysis relatedness, as well as the situation model index causal dependency between sentences. This model successfully predicted the different types of self-explanations as a function of cohesion. We also found that students' prior knowledge interacted particularly with causal dependency.*

RÉSUMÉ. *iSTART (Interactive Strategy Trainer for Active Reading and Thinking) est un tuteur électronique qui entraîne des étudiants à utiliser certaines stratégies de lecture dans le but de comprendre un texte difficile. Les utilisateurs d'iSTART sont amenés à intégrer les*

informations, contenues dans un texte, à une représentation mentale cohérente, par la technique de l'auto-explication. Le but de notre étude était d'examiner comment la cohésion textuelle influençait qualitativement la production de différents types d'auto-explications, particulièrement les stratégies consistant à expliciter des liens entre phrases ou à élaborer le contenu d'un texte. Dans cette optique, nous avons développé un modèle qui calcule deux genres de cohésion textuelle à partir d'indices linguistiques différents contribuant, d'une part, à la construction de la base de texte, à savoir, la répétition de racines de mots et la similarité sémantique entre phrases (Analyse Sémantique Latente), et d'autre part, à la construction d'un modèle de situation, comme la dépendance causale. Ce modèle permet de prédire différents types d'auto-explications en fonction du degré de cohésion textuelle calculée. Nous montrons en outre que les connaissances initiales interagissent particulièrement avec la cohésion situationnelle, calculée à partir de la dépendance causale entre phrases.

KEYWORDS: *Checking algorithms, Constraints satisfaction, Latent Semantic Analysis (LSA), Phase transition, Text dependancy.*

MOTS-CLÉS : *Test d'algorithmes, Satisfactions de contraintes, Analyses de la Sémantique Latente, dépendance textuelle.*

1. Self-explanations in iSTART

The Interactive Strategy Trainer for Active Reading and Thinking (iSTART) is a computational program that provides students with automated training on appropriate reading strategies to use while reading difficult texts (Van Dijk & Kintsch, 1983; McNamara, Levinstein & Boonthum, 2004). iSTART is grounded on the success of Self-Explanation Reading Training (SERT, McNamara, 2004; McNamara & Scott, 1999). It incorporates theories of text comprehension (Kintsch, 1998) and active thinking (Chi, Sotta, & de Leeuw, 1994) to train students on reading strategies that help them understand difficult texts.

Self-explanation refers to the process of making active explanations of text meaning (McNamara, 2004). It is a comprehension monitoring technique that has been shown to improve the understanding of challenging texts (Chi & Bassok, 1989; Chi, de Leeuw, Chiu, & Lavancher, 1994). Students work with iSTART in a three-step sequence, including the introduction phase, demonstration phase, and practice

phase. During the introduction phase, students watch a discussion on self-explanation strategies between artificial agents (a teacher and two students). In the demonstration phase, students are asked to identify and locate strategies used in computer-generated examples of self-explanations. Finally, during the practice phase, students self-explain sentences from texts while attempting to use the reading strategies learned in the previous steps.

During the practice phase, students are asked to write down their self-explanations. Texts are presented sentence-by-sentence. Some sentences are presented in red font, which signals the target sentences to be self-explained. McNamara (2004) described six different reading strategies that the students can use when producing self-explanations. The first is being aware of understanding (i.e., comprehension monitoring). The second is generalizing the content of a text segment or restating it with different words (i.e., paraphrasing). The third is the generation of domain-specific knowledge based inferences (i.e., elaboration). The fourth is the generation of domain-general knowledge based inferences (i.e., logic and common sense). The fifth is making predictions about upcoming text sentence content (i.e., prediction). The sixth is making reference to previous text sentences or stating explicit relations between sentences, particularly between target sentence and previous sentences (i.e., bridging).

McNamara (2004) showed that self-explanation training improved students' comprehension. Specifically, after training, low-knowledge participants were able to form a more consistent textbase representation of the text, in comparison to low-knowledge participants, who did not receive training (as measured by comprehension questions). Interestingly, an increase in logic and common sense strategy use, but not the particular use of the bridging strategy, explained the training benefit on textbase comprehension. Even though the use of the bridging strategy was positively correlated with comprehension, it did not increase after training in comparison to control group. However, use of the bridging strategy was significantly associated with use of the paraphrases strategy and the quality of paraphrases made by low-knowledge readers increased after training. Hence, (1) learning to more actively process text reduces miscomprehensions, despite knowledge deficits, (2) increases

the use of knowledge-based information, and (3) improves paraphrasing of target sentences.

Because self-explanation is an active process, we can interpret McNamara's (2004) results in a more specific way. First, the fact that explanations making references to previous sentences (bridging) was not affected by training, or prior knowledge, could be interpreted as a result of effortless generation of bridging inferences. However, training led to active and better attempts to relate consistent parts of target sentences to previous sentences, assumedly by improving students' paraphrase accuracy. Second, the increase in logic and common sense in explanations after training is likely another consequence of active processing. In comparison to low-knowledge participants, high-knowledge participants generated more domain-specific knowledge inferences. Specifically, participants elaborated text content as a function of available prior knowledge: low-knowledge students used general knowledge and high-knowledge students used more specific, domain knowledge. In summary, we can reasonably hypothesize that self-explanation training led students to actively bridge and elaborate by making them aware of links and gaps within the text. In other words they became aware of the text cohesion.

In this study, we investigate the possibility that linguistic cues can be identified that differentially influence readers' likelihood of generating bridging as compared to elaborative inferences. Moreover, we expect that the cues corresponding to the textbase and situation model representations are differentially used by high-knowledge and low-knowledge readers.

2. Bridging and Elaboration

A large body of research has addressed how text cohesion guides the formation of bridging and elaborative inferences during reading (Gernsbacher, 1990; Kinstch, 1993; Magliano, Zwaan & Graesser, 1999; McKoon & Ratcliff, 1992; McNamara & Kinstch, 1996; Schmalhofer, McDaniel, & Keefe, 2002; Zwaan, Langstone & Graesser, 1995). We can distinguish between inferences that fill in gaps or elaborate the propositional representation of the text (the textbase), and those that

maintain the continuity or enrich the representation of situation described by the text. In both cases, inference processes can be considered knowledge dependent, even though inference generation at a textbase level - involving bridging inferences, associative knowledge elaborations and transitivity inferences - are reputed to be more automatic than logical reasoning or situation model elaboration, which involve a search of relevant knowledge through long-term memory (Kintsch, 1993).

At the textbase level, both bridging and elaborative inferences can be generated. Textbase coherence is mainly driven by the presence of argument or semantic overlap between text segments (Foltz, Kintsch, & Landauer, 1998; Kintsch, & van Dijk, 1978; Myers & O'Brien, 1998). When text segments are sufficiently related, readers are able to make bridging inferences effortlessly; bridging information is readily accessible in the text or in memory, and there is no need to generate elaborative inferences, unless the readers' goal is to strategically make elaborative inferences (McKoon & Ratcliff, 1992). However, when information to scaffold bridging inferences is not readily accessible, knowledge elaborations are still easily accessible. The accessibility is explained by activation of relevant associates in semantic memory (Ericsson & Kintsch, 1995; Kintsch, 1998, Kintsch, 1993). For example, in "A car stopped. The door opened", knowing that a door is a part of car is quite effortless.

The construction of a coherent mental representation of the situation described by the text is more complex than the construction of a coherent textbase representation. Constructing a situation-model representation, readers often need to rely on highly integrated representations involving causal or other functional dimensions to create links between currently processed and previously encoded information (Magliano et al., 1999). If such a situation model representation has not been created by the reader, or if that representation is incoherent, then those links will not be made. In other cases, incoherence may not even be detected, unless processing argument overlap reveals it (Albrecht & Myers, 1995; 1998). For example, Albrecht et al. (1995) showed that inconsistency between a protagonist's goals and results of an action, mentioned at the end of a narrative, was difficult to notice unless text segments involving the results of the action shared an argument with the text segment stating the goal. However, if sufficient relevant knowledge is available,

and/or if readers are engaged in active processing, bridging and elaborative inferences can be made at a situation model level too (Best, Rowe, Ozuru, & McNamara, 2005; Magliano, et al. 1999; Todaro, Magliano, Millis, Kurby, & McNamara, in press).

Zwaan and Radvansky (1998) proposed that readers' situation model representations are driven by five types of relations within the text: temporal, causal, intentional, spatial, and protagonist. The five types of relations have been developed as cohesion cues in Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara, Louwerse, & Graesser, in press). Among the 700 linguistic cues Coh-Metrix usually computes, some are specifically dedicated to measuring the hypothetical incidence of causal, temporal, spatial and intentional relationships through the surface form of the text. In this framework, these types of relations can be referred to as comprising situational cohesion (McNamara & Magliano, in press). In support of this assumption, Magliano et al. (1999) and Todaro, et al. (in press) demonstrated that discontinuities in situational cohesion resulted in readers making more elaborations than bridging inferences in self-explanations.

In the present study, we expect that cues in the text supporting argument overlap as well as sentences similarity to be more related to the reader's textbase representation, whereas cues related to situational cohesion should be more predictive of the reader's situation model construction. Bridging inferences are more likely to be generated in self-explanations when texts are mostly cohesive because bridging information is readily accessible in the text itself or in memory and there is no need to elaborate text's content. Elaborative inferences should be generated in self-explanations when the text is not sufficiently cohesive and when readers need to rely on prior knowledge to self-explain the text.

3. Prior Knowledge, Levels of Understanding, and Cohesion

McNamara, Kintsch, Songer, and Kintsch (1996) provided one demonstration of how text cohesion and prior knowledge interact to influence comprehension. These researchers assessed comprehension at two levels of understanding, textbase and situation model. High-knowledge readers were generally more accurate than low-knowledge

readers on text comprehension assessments. Low-knowledge readers' comprehension was more apparent on textbase measures than on situation model measures, especially when they had read high-cohesion texts. In contrast, high-knowledge readers were able to take advantage of low-cohesion texts. As such, comprehension was more apparent on situation model measures. In general, high-knowledge readers benefited from low-cohesion texts because they were able to elaborate their mental representation of the text at a situation model level; whereas low-knowledge readers performed better on high-cohesion texts predominantly at a textbase level.

O'Reilly and McNamara (2006) further found that the interaction between cohesion and prior knowledge was modulated by reading skills: among high-knowledge participants, those with lower reading skills benefited from low-cohesion texts; in contrast to skilled readers for whom text cohesion had little effect. Moreover, reading skill tended to help low-knowledge readers comprehend high-cohesion texts at a situation model level.

The findings of McNamara et al. (1996) and O'Reilly et al. (2006) are consistent with theories of encoding such as long-term working memory (Ericsson et al., 1995; Kintsch, 1998). Long-term working memory theory predicts that readers are able to encode information by associating it with cues that belong to a mental retrieval structure. At a textbase level, related information in the textbase representation plays the role of retrieval cues. Readers are able to associate information with previously encoded information by means of argument overlap (Kintsch, 1988; Kintsch et al., 1978) or semantic similarity (Foltz, et al., 1998; Shapiro & McNamara, 2000). At a situation model level, when a rupture disturbs the smooth process of text comprehension, skilled and/or knowledgeable readers are able to link the non-related pieces of information in a general representation of the text by means of retrieval structure that results from the generalization, during reading, of the encoded information (Bellissens & Denhière, 2004; Ericsson & Delaney, 1999).

3.1. Hypotheses

The theoretical frameworks and empirical evidence described in the previous sections lead us to formulate the following hypotheses: (i) When text cohesion is relatively high, bridging inferences are more likely to be generated in self-explanations. In contrast, (ii) when textual cohesion is relatively low, elaborative inferences are more likely to be generated. Moreover, based on McNamara et al. (1996) and O'Reilly et al. (2006), we assume that (iii) high and low-knowledge readers are able to understand and explain a text at a textbase level, but high-knowledge readers are more likely to show an advantage and deeply understand at a situation model level as well; hence we expect an interaction between prior knowledge and textual cohesion when the textual cohesion indices include measures of situational cohesion. Finally, (iv) high-knowledge readers should generate more inferences than low-knowledge readers.

3.2. Textual Cohesion

We address these hypotheses in the construction of a computational model that makes a distinction between textbase and situational cohesion. Our goal was to more specifically define text cohesion in order to identify linguistic indices that predict the conditions that lead students to generate bridging inferences as compared to elaborative inferences in self-explanations. To identify textbase and situational cohesion indices present in a text, we turn to a Natural Language Processing tool called Coh-Matrix.

3.2.1. Coh-Matrix and Textual Cohesion

Coh-Matrix is a computational tool that measures more than 700 indices of text cohesion (Graesser et al., 2004; McNamara et al., in press). Coh-Matrix has been used to evaluate if texts are more or less cohesive by means of a variety of cues in the text, such as semantic similarity, word overlap, causality, temporality, and so on. Each type of cue is assessed using various indices that have been developed with Coh-Matrix. For example, Latent Semantic Analysis (Landauer & Dumais, 1997) is used to assess semantic similarity. Some cohesion indices, such as LSA measures and argument overlap, can be associated with textbase com-

prehension (Shapiro et al., 2000). In contrast, causal, temporal, spatial, and intentional measures are associated with situation model comprehension (Zwaan et al., 1998). A text with higher values on a range of Coh-Metrix indices is considered more cohesive, and constitutes a multidimensional and global textual cohesion factor that has impact on reading and comprehension (Duran, Bellissens, Taylor, & McNamara, 2007).

Generally, Coh-Metrix measures global textual cohesion. This is particularly useful if, for instance, one's goal is to compare the cohesion of texts within a corpus. However, our goal is to predict the various types of inference generation in students' self-explanations. According to our hypotheses we need to predict various types of relations within the texts that link a specific sentence to previous ones. As such we want to show that Coh-Metrix can also be useful for measuring cohesion relations between text sentences and not only a general global textual cohesion.

3.2.1.1. Cohesion between text sentences

We assume that text sentences are more or less related to previous ones by means of several relations. We envision two classes of relations, supported by either (1) cues of relatedness between sentences (e.g., argument overlap and semantic similarity) or (2) cues that are supposed to drive the construction of situation model representation (e.g., causal relatedness).

Relatedness between sentences has been investigated in a recent study (Wolfe, Magliano, & Larsen, 2005). Wolfe et al. (2005) manipulated LSA relatedness between sentences, as well as causal relatedness. Their materials were composed of sentence pairs using a prime and a target sentence. They computed one LSA vector for each sentence, as the centroid of sentence word vectors; and LSA sentence relatedness was the cosine between the prime and target sentence vectors. Causal relatedness was based on experimenter intuition and validated in a pilot study. Wolfe et al. (2005) found that the degree of causal relatedness between sentences influenced both processing time and recall of the target sentences. Target sentences causally related to the prime were read faster and better recalled. LSA relatedness also influenced processing time and recall. Target sentences semantically related to the prime were

better recalled and read faster, but only when causal relatedness was low.

First of all, these results suggest that LSA can be used to assess semantic relatedness between sentences as comprehension units (Landauer, McNamara, Dennis, & Kintsch, 2007). Second, Wolfe et al. (2005) explained the interactions between semantic and causal relatedness in terms of bottom-up and top-down processes in the construction of a coherent discourse representation. When the causal relation was hard to establish and semantic relatedness between sentences was high, bottom up processes, such as semantic activation (Kintsch, 1988), could be sufficient to infer causal relations in working memory. In contrast, when the causal relationship was difficult to establish and semantic relatedness was low, the content of a causal inference could not be activated directly from semantic memory and must be constructed. Therefore, in the following, we will distinguish between sentential relationships involving semantic relatedness and relationships involving situational relations such as causality.

3.2.1.2. Relatedness and dependency

Some sentences are considered to be related to preceding sentences because they are semantically similar, or share arguments with preceding sentences; and some other sentences, would be considered as not only sharing similar information but also modifying the redundant information. Modification of redundant information can be categorized into different classes depending on the category of the modification. Redundant information can be modified, for example, by a causal verb or can be localized in space or time. Respectively, we would refer to causal, spatial, or temporal dependencies. Note, however, that although causal relationship implies causal dependency, causal dependency does not imply causal relationship. We use the term *dependency* to mean the *potentiality* to create or make explicit a causal relationship between two pieces of information as a function of semantic relatedness. Indeed, Wolfe et al. (2005) explained how semantic relatedness can play an important role in the creation of a causal relationship by rendering information readily available that is relevant to the causal relationship.

More specifically, when two sentences share similar information, we will refer to relatedness; whereas when two sentences share similar information that is modified, we will refer to text dependency. Thus, in the following, text cohesion will be defined by both relatedness and text dependency.

Consider the following sentences.

- a) “Orlando slept all night in ignorance.”
- b) “He had been kissed by a queen without knowing it.”

(*Orlando*, Virginia Woolf)

Considering sentences (a) and (b), we would say that the relationship is conveyed by the semantic relatedness between “ignorance” and “without knowing”, but also by “Orlando” and “He”. “Ignorance” is modified in sentence (b) by the fact that information like “kissed by a queen” is added. We would say sentences (a) and (b) are related by causal dependency because information is redundant between the two sentences and information is causally modified.

In the next example,

- c) “Terrence needed a medical insurance.”
- d) “He finally found a job, last week.”
- e) “I must say, I am happy.”

there is no relatedness and also no modification of redundant information between sentences (d) and (e). But something needs to be explained. Making a causal bridge such as “The narrator is happy because Terrence found a job” would generate a shallow self-explanation, in comparison to elaboration of the text content as “The narrator is happy because he is relieved that his friend will be able to get medical insurance through his new contract”. In the second self-explanation, causal inferences elaborate the situation described by the text and necessitate the availability of knowledge about job contracts and medical insurance.

3.2.2. *Text Relatedness and Dependency Model*

To operationalize relatedness and text dependency, we constructed networks of text sentences. For a given text, each sentence was repre-

sented by a node in a network, and the connection between each sentence and the previous sentences was represented by the edges of the network. The connection between the sentences was either weighted by relatedness or dependency measures. A different network was built for each target sentence in a text. As depicted in Figure 1, because the target sentences located at the end of the text were by definition preceded by more sentences, the networks corresponding to the target sentences at the end of the text were naturally larger. For each network, the activation of the target sentence was computed by spreading activation on the basis of the integration phase algorithm used in the Construction-Integration model (Kinstch, 1988).

3.2.2.1. Relatedness values

As stated above, we distinguished between relatedness and text dependency. In the present study, relatedness was instantiated by two separate measures. The first one was Coh-Metrix stem overlap proportion measure between sentences and the second one was the LSA cosine between sentences.

The Coh-Metrix stem overlap proportion measure varies between 0 and 1. That measure presents four interesting characteristics: (1) it is a measure of redundancy; (2) it can be computed automatically; (3) the overlap measure is based on word stem and then allows inflected variations as true redundancy; and (4) overlap measure can involve either a predicate or an argument, which increases the probability to capture argument overlap between sentences. In Coh-Metrix, stem overlap proportion can be computed between adjacent sentences or distant sentences. As we needed to compute overlap between all pairs of sentences, we used stem overlap proportion between adjacent sentences for each pair.

The LSA cosine varies between -1 and 1. On top of the four previous characteristics, it can compute semantic similarity between sentences that do not share word stems, but involve synonyms, antonyms, or meronyms. Indeed, words like man and woman are close in LSA space (General-Reading-up-to-1st-year-college, TASA corpus; Touchstone Applied Science Associates, Inc.), hence $\text{cosine}(\text{man} \wedge \text{woman}) = .37$. The same is true for words like hand and finger, $\text{cosine}(\text{hand}$

\wedge finger) = .61. Moreover, LSA is virtually insensitive to negation, in the sense that, in LSA, negating a proposition does not mean that the proposition and its negation cannot be very related, which is an advantage when one wants to compute relatedness; $\text{cosine}(\text{John drinks wine} \wedge \text{John does not drink wine}) = .94$. To illustrate the advantage of this insensitivity to negation, imagine a particular dialogue in which one character asks: “does John drink wine?” and the other character answers: “no! John does not drink wine.” It is impossible to claim that the question and the answer are not related. If a third character comes and says “yes, John does drink wine”, we are still talking about the same thing. Agreement issue between the different characters is not a question of relatedness but a question of discourse comprehension. However, LSA should be combined with a comprehension model in order to optimize its utility in pursuing such text comprehension issues (Kintsch, 1998; Kintsch, Patel, & Ericsson, 1999; Lemaire, Denhière, Bellissens, & Jhean-Larose, 2006).

3.2.2.2. Text dependency values

Text dependency is a complex variable. As stated in the previous section, (1) text dependency means that redundant information between two text segments has been modified by adding causal, temporal, spatial, functional (goal or intention) or attribute information; and (2) after Wolf et al. (2005), we assume that the degree of text dependency is a function of relatedness. As such, we computed a particular text dependency value, causal dependency. Again, we used Coh-Metrix to compute causality.

The original Coh-Metrix causality index, called Causal Link, is computed by summing the proportion of causal verbs and causal particles per 1000 words. This measure is an approximation of the hypothetical incidence of causality relationships in a text. We computed Causal Link between all pairs of sentences in a same text. Then we normalized the causal link measure, and applied the following formula. Causal dependency between two sentences was C:

$$c = \frac{x(S+L)}{2} \quad (1)$$

in which, x is a normalized measure of Causal Link, S is word stem relatedness, and L is LSA relatedness. Any negative value of LSA relatedness is then replaced by 0s. If S and L are equal to 0, C is equal to 0. If x , S and L are all equal to 1, which is their maximum value, C is equal to 1. Hence C varies from 0 to 1.

3.2.2.3. Networks construction

For a given text of n sentences, relatedness or text dependency values were calculated between each pair of sentences. The process involved the construction of several networks of texts. For example, if the text comprised four sentences, we constructed a 2-nodes, 3-nodes and 4-nodes networks, to respectively compute the relatedness or dependency values of the second, the third, and the fourth sentences of the text (see Figure 1).

3.2.2.4. Sentence relatedness and dependency final activation values

Spreading activation in such networks results in obtaining one final activation value for each sentence node. The final activation value of a sentence node in a relatedness network was the sentence relatedness value (here, LSA or word stem relatedness); and the final activation value of a sentence node in a text dependency network was the sentence dependency value (here, causal dependency value).

4. Experiment

4.1. *Participants*

Seventy-four high-school (10th and 11th grade), and fifty-eight college students participated in the experiment. Students were paid for their participation.

4.2. *Materials*

4.2.1. *Text*

Six texts of about 24 sentences and 400 words were selected from high-school textbooks. There were 2 texts per domain and 3 domains:

A Text Relatedness & Dependency Computational Model

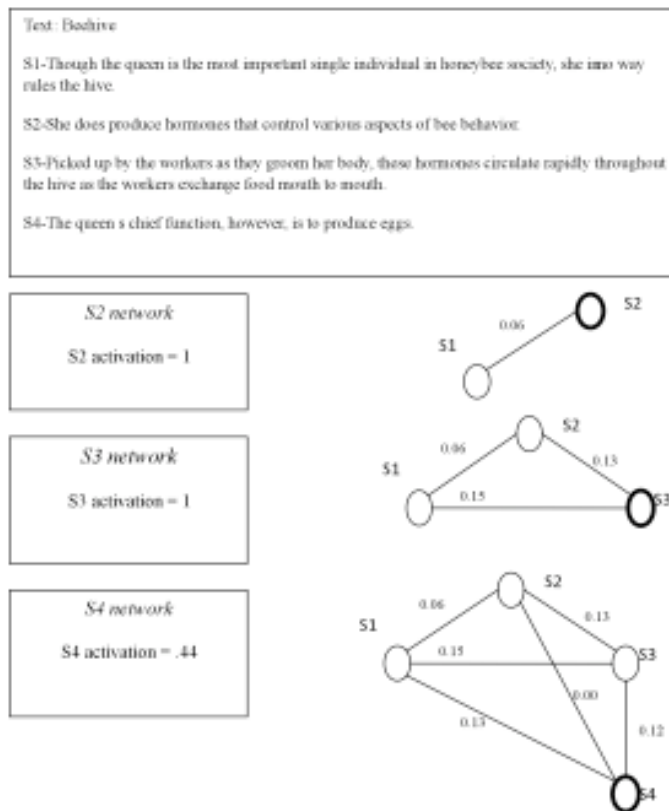


Figure 1: Example of networks construction. Sentences S1, S2, S3 and S4 are linked. Link weights are causal dependency values. S2 dependency activation results from spreading activation through S2 network, S3 dependency activation from S3 network, and so on.

Science (Built for Flight, and Beehive), History: (The New World, and The Americas), and Literature (The Coat, and Invisible Man). The texts were selected with a particular focus on the Flesh-Kincaid grade level index, word frequency, and cohesion as measured by Coh-Metrix. The texts could be qualified as medium cohesive (argument overlap score was $aos = 0.5$; $0 \leq aos \leq 1$), and on 10th Kincaid grade level (out of 12). Considering the populations tested (late high-school and college), we can consider that the texts were of medium difficulty for high-school population and easy for college population.

In order to choose target sentences, two methods were used: the first one involved human expert choice, the second involved *typicality/importance* model (Kintsch, 2002). Human experts' criteria for target sentence selection involved selecting sentences that were: a) important to understanding the text; and b) either information was dense, required an inference (elaborative or bridging) or more extensive elaboration (based on common sense and personal experience).

In terms of *typicality/importance* model, typicality indicated a sentence in a paragraph that was most like all other sentences in the paragraph (Kintsch, 2002). Typicality was calculated by comparing each sentence in the paragraph to every other sentence through LSA. The sentence with the highest LSA score was credited as being the most typical. Importance was calculated by comparing each sentence to the paragraph as a whole. The sentence with the highest LSA score was credited as being the most important.

Finally, one sentence was targeted if it was selected by human experts and appeared to be highly important and typical in the *typicality/importance* model. It resulted that 8 sentences by texts were chosen.

4.2.2. *Prior knowledge test*

Participants were given a multiple choice test of prior knowledge that consisted of 30 questions, with 10 questions per knowledge domain (e.g., science, history, literature). Time to answer was limited to 15 minutes.

Examples of prior knowledge questions:

Science question: The poisons produced by some bacteria are called

- a) antibiotics.
- b) toxins.
- c) pathogens.
- d) oncogenes.

History question: Which statement best describes a characteristic of the Renaissance in Europe?

- a) The social structure became very rigid.
- b) Creativity in the arts was encouraged.
- c) The political structure was similar to that of the Roman Empire.
- d) Humanism decreased in importance.

History question: Which theme is most prominent in the book “Catcher in the Rye?”

- a) teenagers’ identity crisis.
- b) teenagers’ curiosity for adult life.
- c) teenagers’ desire for success.
- d) teenagers’ drug problem.

Prior knowledge score was equal to the number of correct answers divided by the number of questions.

4.2.3. Procedure

The same procedure was used for high-school and college groups. The goal of the procedure was to collect self-explanations generated by participants. We asked two experts to rate the self-explanations.

4.2.3.1. Self-explanation task

The self-explanation task involved self-paced reading of a text, one sentence at a time, on a computer screen (iSTART interface). Sentences were continuously displayed on the screen so participants could reread sentences. For eight times in a text, a target sentence was presented and participants were asked to self-explain the text. The eight target sentences were signaled by red font on the computer screen. Participants had to write down their self-explanations in a box on the screen.

4.2.3.2. Self-explanation ratings

Two experts scored the 6384 collected self-explanations. Their task was to make explicit whether a given self-explanation included information that was not a paraphrase of the given target sentence, and whether added information came from the text itself (i.e., bridging inferences) or not (i.e., elaborative inferences). The coding scheme considered three dimensions: (i) the extent to which a self-explanation overlapped with the target sentence; (ii) the extent to which a self-explanation added information to the text or the target sentence, and (iii) whether the source of any added information was from the text itself or from the assumed reader's prior knowledge. When the information contained in a self-explanation was isolated to the target sentence, it was coded as a *paraphrase* or a *repeat*. When it came from some previous sentences, it was coded it as a *bridging inference*. When the information was not present in the text, it was coded it as an *elaborative inference*. Reliability was established between the raters on the basis of a sample of the self-explanations ($\kappa = 0.67$), then each of the raters coded half of the protocols.

4.2.3.3. Text sentence relatedness and dependency values.

Our text relatedness and dependency model was applied to the six texts used in the experiment. For each text, we computed the two measures of relatedness between all text sentences: (i) word stem overlap; (ii) LSA semantic similarity; and the measure of text dependency: causal dependency (as computed by Equation 1). For each measure, we constructed one networks by target sentence, and spread activation through all networks. It resulted that each target sentence was characterized by three variables: word stem and LSA relatedness final activation values, and one causal dependency final activation value. Over the six texts, distribution of three variables differed noticeably. A two-step clustering algorithm with Akaike Information Criterion (AIC) computed Log-Likelihood distances between data points in the 48 values of each variable.

Within each group of values, the algorithm converged on *different number of clusters* by partitioning the variance so as to maximize the between-cluster variation and minimize the within-cluster variation. Ta-

A Text Relatedness & Dependency Computational Model

bles *1a* and *1b* present relatedness and dependency values of the six text target sentences, and the particular results of the two-step clustering. Tables *1a* and *1b* show that word stem sentence relatedness values have been partitioned into three clusters (*High relatedness*: $M = .93$, $SD = .09$, *Number of sentences* (n) = 15; *Medium relatedness*: $M = .46$, $SD = .13$, $n = 19$; *Low relatedness*: $M = .07$, $SD = .07$, $n = 14$); LSA sentence relatedness values have been partitioned into two clusters (*High relatedness*: $M = .88$, $SD = .10$, $n = 27$; *Low relatedness*: $M = .49$, $SD = .18$, $n = 21$); Finally, causal sentence dependency values have been classified into four clusters (*High dependency*: $M = 1.00$, $SD = .00$, $n = 11$; *Medium high dependency*: $M = .70$, $SD = .05$, $n = 7$; *Medium low dependency*: $M = .44$, $SD = .07$, $n = 17$; *Low dependency*: $M = .13$, $SD = .09$, $n = 13$).

text	sentence	Relatedness values		Dependency values	Relatedness clusters		Dependency clusters
		Word stem	LSA	Causal	Word stem	LSA	Causal
<i>Built for flight</i>	S03	1.00	1.00	1.00	1	1	1
	S08	0.68	0.88	0.35	1	1	3
	S10	0.33	0.60	0.54	2	2	3
	S13	0.33	0.67	0.42	2	2	3
	S16	0.29	0.89	0.76	3	1	2
	S18	0.13	0.14	0.04	3	2	4
	S21	0.74	0.71	0.22	1	1	4
	S26	0.61	0.82	0.46	2	1	3
<i>Beehive</i>	S02	1.00	1.00	1.00	1	1	1
	S05	1.00	1.00	0.73	1	1	2
	S07	1.00	0.92	0.64	1	1	2
	S09	0.45	0.80	0.67	2	1	2
	S13	0.76	0.80	1.00	1	1	1
	S17	0.87	0.82	0.46	1	1	3
	S19	0.58	0.78	0.52	2	1	3
	S20	1.00	1.00	0.70	1	1	2
<i>The new world</i>	S02	1.00	1.00	1.00	1	1	1
	S05	0.35	1.00	1.00	2	1	1
	S09	0.05	0.59	1.00	3	2	1
	S11	0.11	0.60	0.37	3	2	3
	S16	0.60	0.72	0.38	2	1	3
	S20	0.22	0.40	0.77	3	2	2
	S23	0.92	0.93	0.99	1	1	1
S27	0.55	0.86	0.66	2	1	2	

Table 1: Table 1-a : Relatedness and Dependency Values of Built for Flight, Beehive, and The New World Texts, and the Results of the Two-step Clustering.

text	sentence	Relatedness values		Dependency values	Relatedness clusters		Dependency clusters
		Word stem	LSA	Causal	Word stem	LSA	Causal
<i>The Americas</i>	S02	1.00	1.00	1.00	1	1	1
	S04	0.49	0.81	0.53	2	1	3
	S06	0.15	0.38	0.36	3	2	3
	S09	0.52	0.57	0.54	2	2	3
	S13	0.52	1.00	0.56	2	1	3
	S16	0.87	0.55	0.40	1	2	3
	S20	0.06	0.13	0.00	3	2	4
	S23	0.43	0.51	0.20	2	2	4
	<i>The coat</i>	S03	1.00	0.63	1.00	1	2
S04	1.00	1.00	0.41	1	1	3	
S06	0.12	0.89	0.36	3	1	3	
S09	0.33	0.52	1.00	2	2	1	
S13	0.07	0.76	0.21	3	1	4	
S15	0.30	0.59	0.10	3	2	4	
S17	0.65	0.67	0.03	2	2	4	
S18	0.38	0.66	0.07	2	2	4	
<i>Invisible man</i>	S04	0.86	0.93	1.00	1	1	1
	S06	0.00	0.83	0.18	3	1	4
	S07	0.00	0.74	0.11	3	1	4
	S14	0.03	0.30	0.40	3	2	3
	S18	0.04	0.37	0.27	3	2	4
	S21	0.01	0.60	0.16	3	2	4
	S23	0.34	0.11	0.04	2	2	4
	S26	0.02	0.62	0.48	3	2	3

Table 2: Table 1-b : Relatedness and Dependency Values of The Americas, The Coat, and Invisible Man texts, and the Results of the Two-step Clustering.

5. Results

Analyses of variance (ANOVA) of the proportion of inferences generated in self-explanations were separately conducted for high school and college group results. In the following sections, we present analyses for word stem and LSA sentence relatedness, and causal sentence dependency measures

5.1. Word stem relatedness

In this section, we present the results of inference generation in the different conditions of word stem relatedness. The data were analyzed by means of a 2×3 ANOVA, with inference type (bridging vs. elaboration) and word stem relatedness (high vs. medium vs. low) as within-subjects factors.

5.1.1. High school

Figure 2 displays the mean proportions of generated inferences as a function of inference type and word stem relatedness factors. Word stem relatedness had no significant simple effect on the proportion of generated inferences, $F(2, 146) = 2.08, p = .13$. However, the interaction between inference type and word stem relatedness factors was significant, $F(2, 146) = 11.64, p < .01, MSE = .013$. Participants generated more bridging inferences in high relatedness ($M = .35, SD = .23$) than in low relatedness condition ($M = .28, SD = .18$), $t(73) = 4.24, p < .01$. In contrast, they generated more elaborative inferences in low relatedness ($M = .36, SD = .24$) than in high relatedness condition ($M = .32, SD = .21$), $t(73) = 2.38, p < .01$.

5.1.2. College

No significant effect of word stem relatedness was found with college students, $F(2, 110) = 1.02, p = .35$. Table 2 shows that they generated almost the same mean proportions of bridging and elaborative inferences in the three relatedness conditions, $F(2, 110) < 1$.

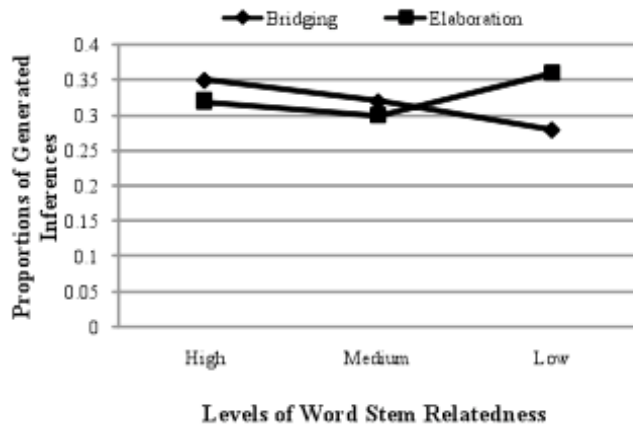


Figure 2: Mean Proportions of Generated Inferences as a Function of Inference Type and Word Stem Relatedness Factors. High School Group

<i>Levels of relatedness</i>	<i>Inference type</i>	
	<i>Bridging</i>	<i>Elaboration</i>
High	0.41 (.23)	0.43 (.22)
Medium	0.43 (.21)	0.44 (.22)
Low	0.44 (.22)	0.42 (.22)

Table 3: Table 2 : Mean Proportions of Generated Inferences as a Function of Inference Type and Word stem Relatedness Factors. College Group.
(standard deviations are into parentheses)

5.2. LSA relatedness

Inference generation in the two conditions of LSA relatedness was analyzed by means of a 2×2 ANOVA, with inference type (bridging vs. elaboration) and LSA relatedness (high vs. low) as within-subjects factors.

5.2.1. High school

Figure 3 presents the mean proportions of inferences generated in self-explanations, by high school students, as a function of LSA relatedness measure. High school students generated significantly more inferences (both bridging and elaboration) on low LSA relatedness target sentences ($M = .33$, $SD = .21$) than on high LSA relatedness target sentences ($M = .31$, $SD = .21$), $F(1, 73) = 6.59$, $p < .05$, $MSE = .006$. The interaction between inference type and relatedness was significant, $F(1, 73) = 6.09$, $p < .05$. The interaction indicated that participants generated more elaborative inferences in low relatedness condition ($M = .35$, $SD = .22$) than in high relatedness condition ($M = .30$, $SD = .20$), $t(73) = 8.19$, $p < .01$. In contrast, the proportion of bridging inferences was the same in low and high relatedness conditions ($M = .32$).

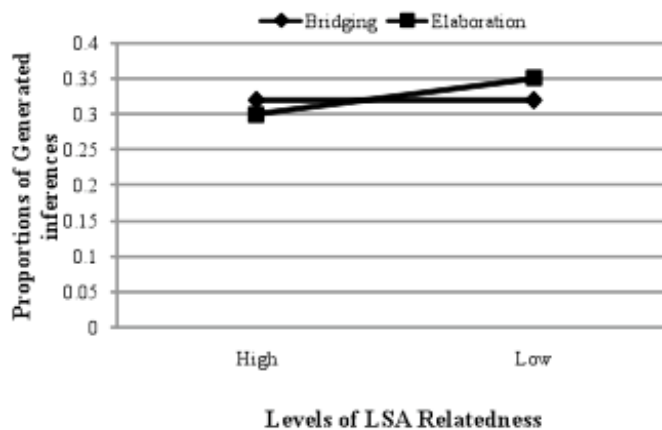


Figure 3: Mean Proportions of Generated Inferences as a Function of Inference Type and LSA Relatedness Factors. High School Group.

5.2.2. College

College students made significantly more inferences in low relatedness condition ($M = .44$, $SD = .21$), than in high relatedness condition ($M = .41$, $SD = .21$), $F(1, 57) = 6.67$, $p < .05$, $MSE = .008$. Interaction between inference type and relatedness was not significant, $F(1, 57) = 1.55$, $p = .22$.

5.3. Causal dependency

Causal dependency factor was made of four conditions: high, medium high, medium low and low dependency. We analyzed the effect of causal dependency on inference generation by means of a 2×4 ANOVA, with inference type and causal dependency as within-subjects factors. Analysis of the data was again conducted separately for high school and college students.

5.3.1. High school

Simple effect of causal dependency factor on the proportion of generated inferences (both bridging and elaboration) was significant, $F(3, 219) = 4.74$, $p < .01$, $MSE = .015$. Participants made more inferences in high dependency ($M = .36$, $SD = .25$) than in low dependency condition ($M = .30$, $SD = .21$), $t(73) = 3.58$, $p < .01$. The interaction between the inference type and causal dependency factors was significant too, $F(3, 219) = 10.58$, $p < .01$, $MSE = .018$. The interaction indicated that participants generated more bridging inferences in high dependency ($M = .39$, $SD = .26$) than in low dependency condition ($M = .26$, $SD = .18$), $t(73) = 5.74$, $p < .01$. In contrast, the proportion of elaborative inferences was not significantly different in high ($M = .35$, $SD = .23$) and low dependency conditions ($M = .32$, $SD = .24$), $t(73) = 1.17$, $p = .24$; nor was it between medium high ($M = .30$, $SD = .24$) and low dependency conditions ($M = .35$, $SD = .23$), $t(73) = 1.89$, $p = .06$, despite the fact the difference seemed larger (see Figure 4).

The results are consistent with the assumption that elaborative inferences are preferably generated in low dependency condition than in high dependency condition, and it was the opposite for bridging inferences.

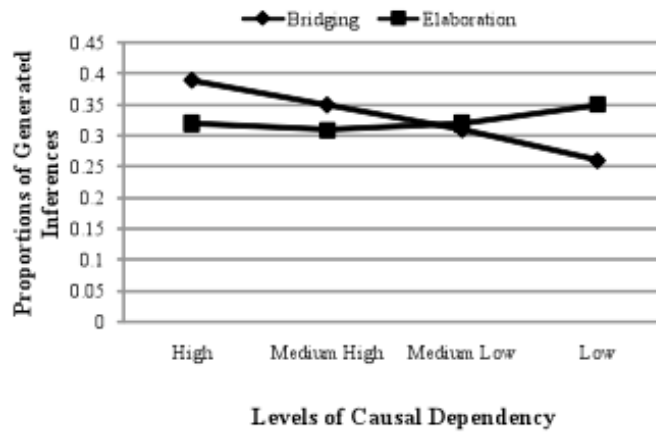


Figure 4: Mean Proportions of Generated Inferences as a Function of Inference Type and Causal Dependency Factors. High School Group.

5.3.2. College

With college group, we found no significant simple effect of causal dependency on inference generation, $F(3, 171) < 1$. However, the interaction between inference type and causal dependency was significant, $F(3, 171) = 4.11, p < .01, MSE = .017$ (see Figure 5). The interaction was explained by the fact that participants generated more bridging inferences in medium high dependency ($M = .46, SD = .26$) than in low dependency condition ($M = .39, SD = .22$), $t(57) = 2.59, p = .01$. In contrast, they generated less elaborative inferences in medium high dependency condition ($M = .40, SD = .24$) than in low dependency ($Mean = .46, SD = .23$), $t(57) = 2.50, p < .05$.

It is interesting to note that the interaction between inference type and dependency was significant for college students, and that was not the case for the interactions when examining the measures of relatedness and inference type.

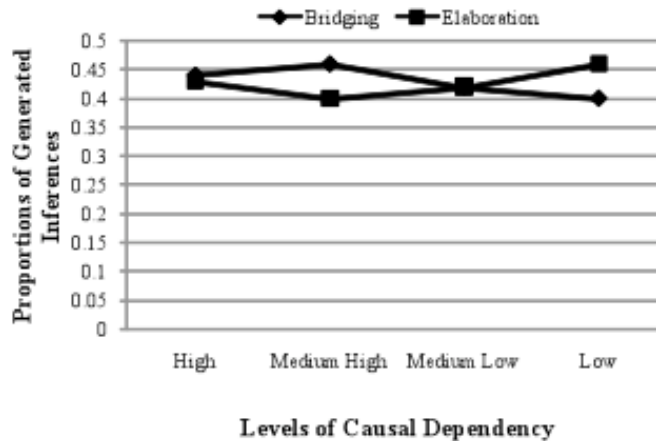


Figure 5: Mean Proportions of Generated Inferences as a Function of Inference Type and Causal Dependency Factors. College Group.

6. Knowledge effects

College students obtained significantly higher prior knowledge scores ($M = .63$, $SD = .16$) than high school students ($M = .54$, $SD = .18$), $F(1, 132) = 7.87$, $p < .01$, $MSE = .031$. To better understand the effect of prior knowledge on the proportions of generated inferences, in each group, we used a mixed model with sentence dependency or sentence relatedness and inference type as within-subjects factors and prior knowledge as a between-subjects factor. Three categories were formed based on the clustering of prior knowledge test scores (high, medium, and low).

6.1. High school

Prior knowledge exerted a simple significant effect on the proportion of generated inferences, $F(2, 71) = 9.05$, $p < .01$, $MSE = .023$. High-knowledge participants generated significantly more inferences ($M = .39$, $SD = .23$) than medium knowledge participants ($M = .30$, $SD = .23$), $F(1, 60) = 4.63$, $p < .05$, $MSE = .023$; and medium knowledge par-

participants generated significantly more inferences than low-knowledge participants ($M = .18$, $SD = .25$), $F(1, 41) = 6.06$, $p < .05$, $MSE = .024$. More knowledgeable high school students generated more inferences.

Prior knowledge did not significantly influence interactions between word stem or LSA relatedness and inference type. In contrast, prior knowledge significantly interacted with causal sentence dependency and inference type. The three-way interaction including causal dependency, inference type and prior knowledge was significant, $F(6, 213) = 2.29$, $p < .05$, $MSE = .017$. Specifically, the interaction between inference type and causal dependency was only significant with high knowledge students, $F(3, 90) = 6.64$, $p < .01$, $MSE = .017$, and medium knowledge students, $F(3, 90) = 7.67$, $p < .01$, $MSE = .017$, but not with low knowledge students, $F(3, 33) < 1$.

Because the number of low knowledge students was small ($n = 12$), in comparison to the number of high ($n = 31$) and medium ($n = 31$) knowledge students, we randomly suppressed the data of 19 high and medium knowledge students. We conducted five different randomizations. Then, for each randomization, we conducted the same analyses. It resulted that the three-way interaction including causal dependency, inference type and prior knowledge was significant with high knowledge students, five times over the five randomizations; with medium knowledge students, only four times; and of course, never with low knowledge students. Hence we decided to further analyze high-knowledge participant results.

Figure 6 displays high-knowledge participants' results. High knowledge participants generated more bridging inferences in high dependency ($M = .45$, $SD = .25$) than in low dependency condition ($M = .31$, $SD = .20$), $t(30) = 5.11$, $p < .01$. In contrast, the proportion of elaborative inferences in high dependency ($M = .39$, $SD = .25$) and low dependency conditions ($M = .41$, $SD = .24$) was not significantly different, $t(30) = .04$; nor it was between medium high ($M = .36$, $SD = .22$) and low dependency conditions ($M = .41$, $SD = .24$), $t(30) = 1.4$, $p = .15$.

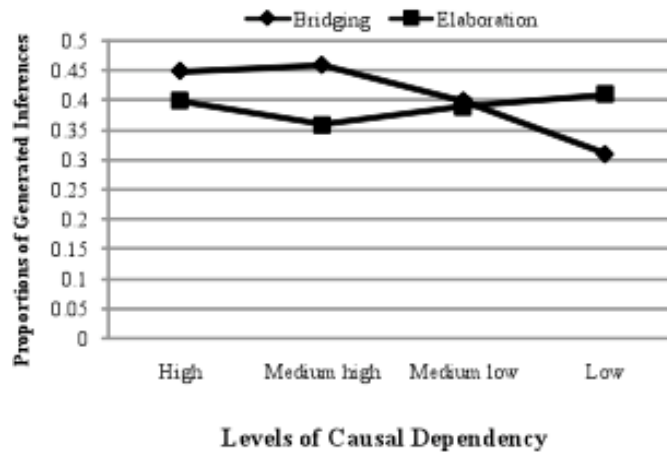


Figure 6: Mean Proportions of Generated Inferences as a Function of Inference Type and Causal Dependency Factors. High-knowledge participants from High School Group.

6.2. College

As in high school group, prior knowledge had a significant effect on the proportion of inferences, college students generated in their self-explanations, $F(2, 55) = 5.77, p < .01, MSE = .020$. High-knowledge students made more inferences ($M = .49, SD = .23$), than medium knowledge students, ($M = .35, SD = .22$), $F(1, 41) = 1.44, p < .01, MSE = .022$; however, we found no significant difference between medium ($M = .35, SD = .22$), and low knowledge students ($M = .38, SD = .23$), $F(1, 27) < 1$. In the college group, however, prior knowledge did not significantly influence the interaction between causal dependency and inference type, $F(6, 165) = 1.92, p = .08$.

7. Discussion

Predicting inference generation as a function of text cohesion and knowledge can be quite useful when the goal is to train students to use particular reading strategies, such as bridging and elaboration. In order

to predict bridging and elaborative inferences that students would generate while self-explaining a text, we constructed a relatedness and text dependency model that automatically determined target sentence relatedness and causal dependency. We made a theoretical and operational distinction between relatedness and dependency. On the one hand, relatedness between sentences included measures of argument overlap or LSA relatedness, which are classically considered as involved at a textbase level of understanding (Foltz et al., 1998; Shapiro et al., 2000). On the other hand, we conceived that relationships between text sentences are not only made of redundant information. The meaning of redundant information is most of the time modified in the subsequent sentence by new information. The modification can belong to different categories. The modification can be causal, temporal, spatial, or functional (goal or intention). The different categories of possible modifications have been classically associated with situation model dimensions (Zwaan et al., 1998). Coh-Metrix can detect markers or different categories of verbs that indicate such categories. Coh-Metrix's Causal Link is such an index. Causal Link measures the incidence of causal verbs and causal particles in a text. When a piece of information is redundant from one sentence to the next one, a particular causal verb or causal particles can modify the content of the redundant information, in the next sentence. We considered that the modification of the redundant information makes the two sentences causally dependent. However, as we stated earlier, causal dependency is not causal relationship but it is a cue for establishing a potential causal relationship.

We assumed that the quality of self-explanations made by students would vary as a function of relatedness and sentence dependency. Self-explaining a related sentence would not necessitate a lot of elaboration because information relevant to bridging was in the text itself. In contrast, weakly related sentences would necessitate greater elaboration by the reader in order for it to be linked with the text. Hence, the more a target sentence was related or dependent to previous sentences, the more bridging inferences would be generated. In contrast, the more a target sentence was not related or dependent, the more elaborative inferences would be used to self-explain the text. We also predicted that high and low-knowledge students that have read moderately low cohesive texts would be able to understand and explain them at a textbase

level, but higher knowledge students would show an advantage for a proximal zone of difficulty and deeply understand the texts at a situation model level too. As such, we expected that only high-knowledge students would be able to use situation model indices, such as causal dependency, to self-explain the texts, and thus we predicted an interaction between prior knowledge and causal dependency.

As should be expected, our results indicate that college students generally have greater knowledge than high school students. These high-knowledge college students generated bridging and elaborative inferences regardless of the level of sentence relatedness. However, they generated more bridging inferences when the target sentence was causally dependent and more elaborative inferences when the target sentence was less dependent. In contrast, high school students were more sensitive to textbase indices, and made more bridging inferences when target sentences were more related to previous sentences, as measured by greater argument overlap and LSA relatedness. They also generated more elaborations when target sentences were less related. Regarding causal dependency, only high school students with greater knowledge generated more bridging inferences for high causal dependency target sentences and more elaborations for less dependent sentences.

In summary, the text relatedness and dependency model has been able to capture a linguistic difference between text sentences that interacts with prior knowledge factors and predicts strategy use in self-explaining a text. We demonstrated that it was possible to automatically predict the type of inferences generated during self-explanations by taking into account LSA relatedness and the causal dependency of target sentences, as well as participants' prior knowledge.

In order to improve such a technique we could compute sentence dependency in terms of other situation model dimensions (e.g., temporal, spatial) besides causality. We know that beginners or experts in specific knowledge domains organize encoded information differently in memory (Caillies, Denhière, & Jhean-Larose, 1999; Caillies, Denhière, & Kintsch, 2002; Denhière & Baudet, 1992; Jhean-Larose, 1991); experts prefer functional relationships whereas beginners prefer temporal/causal relationships. A relatedness and sentence dependency tool should be useful in calculating different kinds of sentence dependency.

We could determine, a priori, which indices of text cohesion and text organization interact best with students' prior knowledge.

A subsidiary issue is the way we computed relatedness. We chose to compute two measures, one based on the Coh-Matrix word stem index and the other using the up-to-12th grade LSA space. A theoretical comparison of the two indices leads us to question the role of knowledge in the perception of relatedness between sentences. It seems that word stem and LSA relatedness were quite similar in the way they influenced self-explanation quality. However, there is no reason to think that another LSA space would have led to the same results, particularly if the space was built from a specific knowledge domain corpus or a corpus made of documents for a grade higher than the 12th grade. This could be a crucial issue if we consider crossing knowledge factors with relatedness or dependency factors. It would likely be useful to test the effects of using different LSA spaces on the determination of relatedness and even dependency.

Finally, in this study, we focused on bridging and elaboration self-explanation strategies. We could have studied the combination of other strategies, in particular the combination of paraphrasing strategies with bridging or elaboration strategies. Combination of different self-explanation strategies is particularly instructive about the way students organize information from a text at different levels of understanding (McNamara, 2004). Future research should further investigate relationship between combination of strategies and relatedness / dependency. Predicting how students are able to use relatedness or dependency indices to self-explain a text, as well as knowing the level of understanding they can use to fully explain a text, could be key information that helps teachers, researchers, and developers of tutoring technologies improve diagnostics and remediation.

References

Albrecht, J. E., & Myers, J. L. (1995). – The role of context in the accessibility of goals during reading.– *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1459-1468.;

Albrecht, J. E., & Myers, J. L. (1998). – Accessing distant text information during reading: Effects of contextual cues. – *Discourse Processes*, 26, 87-107;

Bellissens, C., & Denhière, G. (2004). – Retrieval structure construction during reading: Experimentation & simulation.– *In Proceedings of the Twenty Sixth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.;

Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). – Deep-level comprehension of science texts: The role of the reader and the text. – *Topics in Language Disorders*, 25, 65-83.;

Caillies, S., Denhière, G., & Jhean-Larose, S. (1999). –The intermediate effect : Interaction between prior knowledge and text structure. – *In H. van Oosterdorp S. Goldman (Eds.), The construction of mental representation during reading*, Hillsdale, NJ: Lawrence Erlbaum Associates.;

Caillies, S., Denhière, G., & Kintsch, W. (2002). – The effect of prior knowledge on understanding from text: Evidence from primed recognition. – *European Journal of Cognitive Psychology*, 14, 267-286.;

Chi, M. T. H., & Bassok, M. (1989). – Learning from examples via self-explanations. – *In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser*, Hillsdale, NJ: Lawrence Erlbaum Associates.;

Chi, M. T. H., de Leeuw, N., Chiu, M., LaVancher, C. (1994). – Eliciting self-explanations improves understanding. – *Cognitive Science*, 18, 439-477.;

Chi, M. T. H., Sotta, J. D., de Leeuw, N. (1994) – From things to processes: A theory of conceptual change for learning science concepts. – *Learning and Instruction*, 4, 27-43.;

Duran, N., Bellissens, C., Taylor, R., & McNamara, D.S. (1992). – Quantifying text difficulty with automated indices of cohesion and se-

mantics. – *In Proceedings of the 29th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.;

Ericsson, K. A., & Delaney, P. F. (1999). – Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. – *In A. Miyake P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control*, Cambridge, UK: Cambridge University Press.;

Ericsson, K. A., & Kintsch, W. (1995). – Long-term working memory. – *Psychological Review*, 102, 211-245.;

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998) – The measurement of textual coherence with Latent Semantic Analysis. – *Discourse Processes*, 25, 285-308.;

Gernsbacher, M. A. (1990). – Language comprehension as structure building. – *Hillsdale, NJ: Lawrence Erlbaum Associates.*;

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). – Coh-Metrix: Analysis of text on cohesion and language. – *Behavioral Research Methods, Instruments and Computers*, 36, 193-202.;

Hutchins, J. (1977). – On the structure of scientific texts. – *UEA Papers in Linguistics*, 5, 18-39.;

Jhean-Larose, S. (1991). – L'apprentissage d'un système fonctionnel complexe. – *Psychologie Francaise*, 36, 167-177.;

Kintsch W. (1988). – The role of knowledge in discourse comprehension: a constructive-integration model. – *Psychological Review*, 95, 163-82.;

Kintsch W. (1993). – Information accretion and reduction in text processing: Inferences. – *Discourse Processes*, 16, 193–202.;

Kintsch W. (1998). – Comprehension: A paradigm for cognition. – *New York: Cambridge University Press.*;

Kintsch W. (2002). – On the notion of theme and topic in psychological process model of text comprehension. – In *M. Louwerse, W. Van Peer (Eds.), Thematics: Interdisciplinary Studies*. Amsterdam: Benjamins;

Kintsch W, Patel, V. L., & Ericsson, K. A. (1999). – The role of long-term working memory in text comprehension. – *Psychologia*, 42, 186-198.;

Kintsch W, & Van Dijk, T. A. (1978). – Towards a model of text comprehension and production. – *Psychological Review*, 85, 363-394.;

Landauer, T. K., & Dumais, S. T. (1997). – A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. – *Psychological Review*, 104, 211-240.;

Landauer, T., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). – Handbook of Latent Semantic Analysis, Mahwah, NJ: Lawrence Erlbaum Associates.

Lemaire, B., Denhière, G., Bellissens, C., & Jhean-Larose, S. (2006). – A computational model for simulating text comprehension. – *Behavior Research Methods, Instruments, & Computers*, 38, 628-637.;

Magliano, J. P., Zwaan, R. A., & Graesser, A. (1999). – The role of situational continuity in narrative understanding. – In *H. van Oostendorp & S. R. Goldman (eds.), The construction of mental representations during reading*, Mahwah, NJ: Lawrence Erlbaum Associates.;

McKoon, G., & Ratcliff, R. (1992). – Inference during reading. – *Psychological Review*, 99, 440-466.;

McNamara, D. S (2004). – SERT: Self-explanation reading training. – *Discourse Processes*, 38, 1-30.;

McNamara, D. S., & Kintsch, W. (1996). – Learning from texts: Effects of prior knowledge and text coherence. – *Discourse Processes*, 22, 247-288.;

- McNamara, D.S., & Magliano, J. P.** (in press), – Towards a comprehensive model of comprehension.– *In B. Ross (Ed), The psychology of learning and motivation, vol. 51. New York, NY, US: Elsevier Science.*;
- McNamara, D.S., & Scott, J. L.** (1999). – Training reading strategies.– *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum Associates.*;
- McNamara, D.S., Kintsch, E., Songer, N. B., & Kintsch, W.** (1996). – Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text.– *Cognition and Instruction, 14, 1-43.*;
- McNamara, D.S., Levinstein, I. B., & Boonthum, C** (2004).– iS-TART: Interactive strategy training for active reading and thinking.– *Behavior Research Methods, Instruments, & Computers, 36, 222-233.*;
- Myers, J. L., & O'Brien, E. J.** (1998). – Accessing the discourse representation during reading.– *Discourse Processes, 26, 87-107.*;
- O'Reilly, T., & McNamara, D. S.** (2006). – Reversing the Reverse Cohesion Effect: Good texts can be better for strategic, high-knowledge readers.– *Discourse Processes, 43, 121-152.*;
- Shapiro, A. M., & McNamara, D. S.** (2000).– The use of latent semantic analysis as a tool for the quantitative assessment of understanding and knowledge.– *Journal of Educational Computing Research, 22, 1-36.*;
- Schmalhofer, F., McDaniel, M. A., & Keefe, D.** (2002). – A unified model for predictive and bridging inferences.– *Discourse Processes, 33, 105-132.*;
- Todaro, S. A., Magliano, J. P., Millis, K. K., Kurby, C. A., & McNamara, D. S.**, (in press).– Availability of Information Sources during Reading. *Discourse Processes.*;
- Van Dijk, T. A., & Kintsch, W.** (1983). – Strategies of discourse comprehension.– *New York: Academic Press.*;

Wolfe, B. W., Magliano, J. P., & Larsen, B. (2005). – Causal and semantic relatedness in discourse understanding and representation.– *Discourse Processes*, 39, 165-187.;

Zwaan, R. A, Langston, M. C., & Graesser, A. C. (1995). – The construction of situation models in narrative comprehension: An event-indexing model.– *Psychological Science*, 6, 292-297.;

Zwaan, R. A., & Radvansky G. A. (1998). – Situation models in language comprehension and memory.– *Psychological Bulletin*, 123, 162-185.;

8. Acknowledgements

The research was supported by Institute for Educational Sciences (IES R305G020018-02; R305G040046, R305A080589) and National Science Foundation (NSF REC0241144; IIS-0735682) awards to the 4th author. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the IES or the NSF. We would like to thank all the CSEP Lab members who participated in the collection of the data, and particularly Yasuhiro Ozuru.