# Knowledge Activation in Story Comprehension

Irene-Anna N. Diakidoy[1], Loizos Michael[2], Antonis Kakas[3]

*Department of Psychology, University of Cyprus[1]*
*School of Pure and Applied Sciences, Open University of Cyprus[2]*
*Department of Computer Science, University of Cyprus[3]*

The purpose of this study was to provide behavioral data regarding the world knowledge that readers activate during story comprehension with a view to inform the development of formal and computational representation of knowledge for automated comprehension. The focus of the study was on the amount, the structural aspects, and the coherence function of the activated knowledge. In a knowledge elicitation task, undergraduate students recorded the knowledge they deemed necessary for understanding each sentence in a set of four short stories in the form of property and causal knowledge rules. Results indicated the activation of a limited amount of knowledge that maintained both situational and conceptual connections with input. Situational inferencing and continuity (i.e., carrying over of previously activated ideas) were positively associated and varied as a function of story part. Findings are discussed in relation to dominant theoretical models of narrative comprehension and for their implications regarding the computational representation of knowledge in automated story comprehension.

**Keywords:** *story comprehension; automated comprehension; knowledge activation; inferences*

## 1. Introduction

Text comprehension, and, more specifically, story comprehension rests on utilizing background knowledge about the world that people amass over

***Journal of Cognitive Science*** **18-4: 439-471, 2017**
Date submitted: 10/22/17 Date reviewed: 12/21/17
Date confirmed for publication: 01/08/18

the years through their experience. Several works have exposed this need of world knowledge activation during the process of comprehension aiming to formulate and develop concrete models of representation and use of knowledge for comprehension (Graesser, Millis, & Zwaan, 1997; Kintsch, 1988; Zwaan, 2016). In particular, the task of formulating computational models for story comprehension requires a detailed understanding of the nature of world knowledge in relation to its use for forming, updating and revising a comprehension model. Indeed, simulating human inferential comprehension outcomes rests on representing background knowledge in a computable form that supports mechanisms for representational coherence, elaboration, and revision (Budiu & Anderson, 2004; Frank, Koppen, Noordman, & Vonk, 2003, Kintsch & Mangalath, 2011).

Such computational work has focused primarily on simulating comprehension processes and outcomes in the service of theoretical model testing (e.g., Kintsch, 1998; Singer & Kintsch, 2001; Trabasso & Wiley, 2005; Tzeng, van den Broek, Kendeou, & Lee, 2005). Although informative and productive, this approach has also shown the task to be more complicated than what one may have hoped for, resulting in approaches that end up, among others, ignoring the role of background knowledge, or assuming the presence of only what knowledge is necessary to make the model work as intended (see also Frank, Koppen, Noordman, & Vonk, 2003, 2008). A more promising approach involves the extraction of knowledge from large text corpora on the basis of word co-occurrences (Landauer & Dumais, 1997). However, the questions in this case concern the extent to which lexical or even propositional meanings thus extracted incorporate and/or correspond to the knowledge that people extract from their experiences in the world and which is available to them for text comprehension purposes (Sanford & Garrod, 1998).

In our work, we focus directly on the content and the structure of the underlying knowledge that makes human text comprehension possible and would thus help in automated comprehension. In contrast to dominant approaches, our method is to extract and describe the knowledge that readers use to understand text, in a computational representation form in order to allow direct comparisons between behavioral and computational knowledge input and comprehension outcome. As a result, any gaps or



divergences can be further examined as to their implications regarding the representational characteristics of the knowledge base accessed by readers and the computational modeling of comprehension mechanisms, such as coherence building and revision. Within this context, the specific aim of the present study is to examine the amount, the structural nature, and the coherence function of the knowledge that readers activate in order to understand a set of simple short stories with a common underlying theme.

### 1.1. The Computational Framework

The computational framework for story comprehension in which our study is placed is based on a rule-based knowledge representation and an argumentation-based underlying inferential process. This model is implemented into an automated system, called STAR, standing for "Story Comprehension through Argumentation" (Diakidoy, Kakas, Michael, & Miller, 2014), which forms a comprehension model from the text and has the ability to select appropriate responses to inferential questions and to revise these responses as a story progresses. This system is available online, along with benchmark stories and other material, at: http://cognition.ouc.ac.cy/star.

Comprehension in the STAR framework proceeds by applying a suitable form of argumentative reasoning, as adopted from argumentation theory in AI (e.g., Bench-Capon & Dunne 2007; Rahwan & Simari, 2009), and in line with psychological evidence on the central role of argumentation in human reasoning (Mercier & Sperber, 2011), to integrate the explicit information from a story with the background knowledge assumed to be available to a typical human reader. Knowledge is represented in the form of argument rules that associate a premise (or a set of such) with a position supported by the argument. Argument rules are distinguished in terms of the type of relation they represent into two categories: causal argument rules, (x CAUSES y) or (x STOPS y), which associate causes and effects; property argument rules, (IF x THEN y), which associate properties of objects in the world with another property. It must be noted, however, that property argument rules may also implicate causal relations, especially when the second part of the rule can represent a cause that is offered as an explanation of the first part.

Causal knowledge is of paramount importance for story comprehension. Our computational framework includes a treatment of time and causality grounded on central theories of AI developed initially by McCarthy (1963) with the Situation Calculus, and then by Kowalski and Sergot (1986) with the Event Calculus. This accommodates for the persistence of situations and concepts across time, and their interaction with inferences from causal and property rules. Conflicts in inferences (pertaining to the same time instance) are resolved through implicit general priorities between arguments and other human biases indicating which arguments are acceptable and which ones are defeated in the presence of stronger ones.

Our study then of the nature of world knowledge used by readers was carried out within this context of argumentative causal and property representation and reasoning. A key issue is the selection and the form of the background knowledge to be represented. While well-developed theories of comprehension carry implications regarding the extent, the structure, and the function of the knowledge that human readers activate during story comprehension, these implications are often vague and/or conflicting (see, for example, McNamara & Magliano, 2009). On the other hand, a priori subjective or outcome-based selection of this knowledge obscures the flexibility of human cognition in processing text and generating inferences that give rise to and, simultaneously, reveal mental representations of varying levels of specificity, coherence, and elaboration (Smith & Hancox, 2001). Moreover, this approach would prevent future comparisons between human and automated inferential input and outcomes, confounding, eventually, the evaluation of our computational comprehension framework. Therefore, we have turned to human readers for feedback regarding what knowledge, in what form, and to what extent is activated in a story comprehension context.

### 1.2. Comprehension and Background Knowledge

The outcome of comprehension is the construction of a coherent and sufficiently elaborated mental representation of the text read. The degree of the representational coherence and elaboration that should be achieved cannot be further specified up front, as it can naturally vary as a function of the reader, the text, and the situational context of reading (Graesser, Singer, & Trabasso, 1994; Smith & Hancox, 2001; van den Broek, Lorch, Linderholm, & Gustafson, 2001). The activation of relevant background knowledge is critical for the generation of inferences that supply or specify missing or implicit connections between text parts and elaborate text ideas (Graesser et al., 1997; McNamara & Magliano, 2009). In the case of stories, their typical content ensures that most readers can rely on vast amounts of world knowledge – including knowledge about human motives and actions, causes and consequences, and possible spatial relations and temporal sequences (e.g., Graesser et al., 1997; Suh & Trabasso, 1993; Zwaan, Langston, & Graesser, 1995) – for inference generation. But how much of all that potentially useful knowledge is actually activated and used in the service of story comprehension? Then, considering (a) the cognitive limitations, (b) the kinds of relations that can exist between a given text input and aspects of knowledge, and (c) the representational coherence for a successful outcome, one may also ask in what structural form (or in what way related to input) must the activated knowledge be and to what extent does it function to increase coherence. These issues form the basis for the questions that we investigate in this study.



#### *1.2.1. Amount of activated knowledge.*

Most theories of text comprehension acknowledge cognitive limitations indirectly by postulating fast and automatic memory retrieval mechanisms via spreading activation or resonance, regardless of the relevance of the retrieved ideas in the context or the situation described in the text (e.g., Cook, Halleran, & O'Brien, 1998; Myers & O'Brien, 1998), followed by similarly automatic processes of integration. Memory retrieval mechanisms contribute to the construction of a network of text and knowledge-derived propositions (Kintsch, 1988). Integration is conceptualized as a constraint-satisfaction mechanism that ensures that all but a few (previously activated) text- and knowledge-based ideas are suppressed or dropped from the focus of attention at any given point during reading (Kintsch, 1988; Long & Lea, 2005).

Although a distinction between the two mechanisms carries theoretical importance, focusing computationally on such a fine-grained level becomes quickly unwieldy, as it necessitates the activation of numerous, ultimately suppressed, parts of the knowledge, that are less likely to contribute to the inferential comprehension outcome. Focusing, instead, on the episodic memory content, as it changes across sentences, enables the harvesting of knowledge that readers use to comprehend a story. Specifying the amount of that knowledge without compromising the comprehension outcome is important for computational effectiveness and efficiency. Our first question, then, is: How much knowledge is consciously activated in episodic memory during the processing of each text segment? (Q1)

*1.2.2. Structure of activated knowledge.*

Memory-based comprehension theories that postulate the use of resonance or spreading activation mechanisms operate largely on the assumption of symbolic or sub-symbolic associative network structures of knowledge, with concepts and propositions being the basic units in such structures (Kintsch, 1998; Tzeng et al., 2005). Integration processes prune weakly activated concepts and propositions and pool together or subsume the rest into more general overarching units (i.e., macro-propositions, Kintsch, 1988). The implication of this account is the activation of relatively fine-grained knowledge conceptually related to input regardless of possible contexts of occurrence beyond the linguistic one (e.g., Kintsch & Mangalath, 2011).

Stories, however, are about situations and events within which the meaning of any proposition must be grounded for the story to be understood. Several theoretical models take that into account by postulating either a scenario-based interpretation of propositional content (Sanford & Garrod, 1998) or a situational or event representation in place of a propositional one (Gernsbacher & Robertson, 1992; Kurby & Zacks, 2012; Zwaan, 2016). These accounts maintain some similarity to earlier accounts of knowledge organized in schema-like structures (Schank & Abelson, 1977), and Kintsch (1998) allows for the possibility of the content of nodes to range from concepts to propositions to schemas, with schema-like structures functioning as controls (as opposed to filters) in comprehension. The common underlying implication in these accounts is the existence of multiple situational and functional connections – in addition to conceptual ones – between text input and the knowledge activated (e.g., Norman, 1982).



From our perspective, the "content of the node" notion (or what the node represents) is critical and in need for further specification as it carries different implications regarding both the grain-size of the knowledge and the kinds of its connections to the input that must be computationally represented. Therefore, our starting point is the general assumption that any text segment can activate knowledge that is either conceptually or situationally related to it, whose nature we wish to identify and measure. Our second question, then, is: Do readers rely primarily on general conceptual associations that hold across situations and contexts or do they evoke specific situations within which conceptual associations apply? (Q2a)

However, stories are all about changes from initial states to other modified states (Toolan, 2006), and it is entirely possible that the primary nature of the relations in Q2a may vary in the course of story reading and understanding. According to a scenario-mapping hypothesis (Sanford & Garrod, 1998), scenarios representing situation-specific knowledge are activated passively at the earliest opportunity and drive the mapping of text onto world knowledge. According to Gernsbacher and Robertson (1992), when readers start processing a story they also start laying a foundation upon which subsequent story information will be mapped, while Zwaan et al. (1995) have argued that readers continuously monitor for shifts or breaks in important story dimensions such as time, location, causality, character and intentions. In light of these proposals and supporting evidence (e.g., Dery & Koenig, 2015; Radvansky & Copeland, 2010; Zwaan, 1996), our second question is adapted to examine potential variability as a function of the part of the story that is the current focus of processing: Does the degree of situational specificity of activated knowledge vary as a function of a text segment's position in the story and whether it represents a shift in the story line? (Q2b)

*1.2.3. Function of activated knowledge.*

Shifts and breaks in a story line represent potential coherence breaks that readers must repair in order to end up with a coherent mental representation (or situation model) of the whole story. Although coherence building is naturally a central issue in all psychological theories of comprehension, differences exist regarding the hypothesized mechanisms underlying it. Bottom-up theories postulate construction and integration processes that result in the building of microstructures of connected text-based propositions, and macrostructures containing only highly-activated and re-activated previous input in some more general form (e.g., Kintsch, 1988). Therefore, local coherence involves connecting propositions that are proximal to each other and/or share arguments, while inhibition and generalization seem to underlie global coherence. In contrast, theories that have focused on narrative comprehension emphasize the importance of underlying causal structures in stories as the basis for inferences that connect text segments referring to goals and obstacles with those referring to actions and outcomes (e.g., Graesser et al., 1994; Suh & Trabasso, 1993; Trabasso & van den Broek, 1985; Zwaan et al., 1995).

From a computational perspective, coherence building represents a challenge as it entails the dynamic updating of mental representations with each new textual input, while simultaneously contributing fundamentally to the sought computational outcome of inferential comprehension. Whatever knowledge is activated to support comprehension, a fundamental question concerns its persistence and contribution to coherence. Viewing coherence building in terms of keeping a level of continuity between the updates of the representation may offer a potentially productive initial way to start addressing this computational problem. According to Zwaan et al. (1995), during breaks and shifts in the story line readers must establish a new mental model or alter a previous one by maintaining some and deactivating other ideas included in a previous mental model. The notion of carrying over is computationally feasible, provided we specify what exactly is being carried over.

Although our primary interest lies in the activated knowledge and its coherence function, we have to take into account that what comprises a mental representation at any given point during text processing is a mixture of knowledge and text data, some of which has been carried over from previously constructed representations. At this initial stage, we do not differentiate between previous text input and previously generated inferences, opting, instead, to consider the extent to which continuity (carrying over) of both kinds of elements is maintained during reading and how it may vary in response to input as reading progresses along a



story. Therefore, our third question addresses the coherence function of knowledge indirectly: To what extent does activated knowledge involve the carrying-over of ideas (inferred or text-based) from previous processing (Q3a), and how may that vary as a function of the part of the story (beginning, middle, or ending) that is the focus of current processing? (Q3b)

### 1.4. The Present Study

Our study was designed to help specify to a more satisfactory degree the knowledge needed for an automated system to comprehend a set of four short stories. The stories (7-8 sentences long, preceded by a title) shared the same initial setting and event (the home, doorbell ringing) but had different protagonists and story outcomes. All stories were written to include a turn of events that would constitute a surprise given any expectation that readers may have formed from initial story information. The goal was to harvest knowledge used in the comprehension of stories (as opposed to simple event narrations) and potentially useful across a range of story situations within a context of experience (generalizability issue). To that purpose, we presented each story, sentence by sentence, and asked readers to record the knowledge they deemed necessary for understanding that sentence.

Moreover, the readers in our study were asked to record their activated knowledge in the form of property and/or causal association rules. The selection of this knowledge format was based on considerations of psychological validity (Anderson, 1993; Budiu & Anderson, 2004), and relative compatibility with both situation-specific and general conceptual associations as well as the causal structure of stories (Schunn & Klahr, 1999; Trabasso & Wiley, 2005). Knowledge in this format is also computationally useful as it allows the direct translation of cognitive data into computational data readable by our automated STAR system. To control for any comprehension failures and to ultimately validate the translation and the system, each story was followed by three inferential multiple-choice questions.

With respect to the activated knowledge, we were interested in specifying its amount (Q1), its structural characteristics (Q2), and its coherence-related function (Q3). Given our format of knowledge recording, we operationally defined amount of knowledge activated as the number of rules provided by readers overall, across stories and sentences within stories, and on the average for each sentence read. To account for the possibility of rules manifesting the coordination of multiple aspects of knowledge, we provided our readers also the option of recording their knowledge in rules of varying complexity (single vs. multiple premises per rule). Since multiple premises can indicate the activation of larger amounts of knowledge, we also measured the number of complex rules provided overall and for each sentence in each story as a secondary measure.

To address Q2, we focused on the kind of the relation reflected in the connections between the knowledge and the story and the extent to which it is shown to hold in the context of a particular situation or in general. For example, upon reading the sentence The doorbell rang. one may activate knowledge about humans causing the ring and/or about doorbells being outside and next to doors that hold regardless of context. But one may also activate more context-specific knowledge about a visitor having arrived, or a host running to respond. The "scenario" of a visit reflects a larger chunk of functionally interrelated knowledge that explains the event and allows for multiple but more specific agents (e.g., visitors and hosts, instead of just persons ringing bells and opening doors), actions, and consequences. In order to answer Q2, we compared the number of situation-specific inferences manifested in rules to those reflecting general conceptual associations. Moreover, in light of theoretical and empirical considerations (Gernsbacher & Robertson, 1992; Radvansky & Copeland, 2010; Zwaan, 1996), we examined the possibility of readers activating more situation-specific knowledge during the processing of the middle part of a story in comparison to the beginning part (the setting), or the final part that offered an unexpected ending to the story.

With respect to Q3 regarding the coherence function of knowledge, we conceptualized any continuity of text-based and knowledge-based ideas across the processing of successive story segments as attempts to establish coherence. Continuity is exhibited when a previously encountered or activated idea is brought to bear (carried-over) in the knowledge activated for the processing of the current segment. Therefore, in order to answer Q3, we measured the number of continuous rules, that is, non-redundant rules that made explicit reference to previous text ideas and inferences, and



in relation to non-continuous rules, overall and as a function of story part being processed (as for Q2). However, we were also interested in whether more ideas are carried over during the processing of the middle parts of the stories when compared to the processing of the final parts where readers may be updating situation models in response to unexpected events. Finally, we also reasoned that the extent to which coherence building derives from text-based conceptual integration or scenario-mapping processes may be reflected in the magnitude of the association between continuity and the structural characteristics of the activated knowledge. Therefore, we also examined the extent to which continuity is associated more with situational inferences than conceptual inferences.

## 2. Method

### 2.1. Participants

The sample included 68 undergraduate students majoring in Computer Science (n = 33) and Psychology (n = 35) from a medium-sized state university. There were 22 males and 46 females. The majority of the students (n = 39) were in their first year of study, while the rest (n = 29) were in their third year of study. Study participation was either part of the requirements of an introductory computer science course or for extra credit in a psychology course.

### 2.2. Pre-reading Measures

A vocabulary test and a story recall task were included in order to account for any pre-existing differences in verbal ability that could influence knowledge elicitation and story comprehension outcomes. The vocabulary test was a 50-item synonym test constructed to be used with Greek-speaking young educated adults (university students). The test includes low-frequency words (20 nouns, 10 verbs, 20 adjectives) taken from literary texts not included in high-school curricula. The test's reliability with the present sample was moderate (Cronbach's $\alpha$= .74). The story recall task required students to read and recall two short stories, 163- and 249-words long, respectively (cf. Diakidoy et al., 2014). Each story followed the typical structure of setting plus single episode and included elements of surprise.

Students read both stories in counter-balanced order and, after a short filler task, were asked to write down all they could remember from each story. Students' recall protocols were transcribed and parsed into clauses by two independent judges (92 % agreement, Cohen's K = .84, $p < .01$). Subsequently, each recall clause was scored as to whether it represented verbatim or paraphrase recall of a text clause (91 %, agreement, Cohen's K = .82, $p < .01$) yielding an overall Story Recall Score (proportion of story clauses recalled) for each student.

**2.3. Experimental Texts and Inferential Questions**

Four short stories with titles were written for the purposes of the study. All stories were 50 to 92 words long and included seven to eight sentences. The first sentence was "The doorbell rang." and was the same across stories. The content of each story, however, diverged after the first sentence to present different scenarios with one unexpected event or situation (see Appendix). Each story included three parts. The initial part presented the common initiating event and the protagonist. The next two to three sentences comprised the middle part of the story and presented more information regarding the situation and the protagonist's motives, needs, or expectations. The final two to three sentences (the final part) presented information about protagonists, actions, or outcomes that could be considered unexpected given the information presented in the middle part of the story.

Each story was followed by three inferential comprehension questions in order to ensure comprehension and to provide a baseline for (eventually) examining the performance of the automated story comprehension system STAR. Each question required the reader to either infer causes and/or consequences of explicit story events or to determine the identity of story characters. The format of the questions was multiple-choice with only one of the three choices supported by the story. Nevertheless, neither the correct choice nor any of the distractors were explicitly mentioned in the texts (see Appendix).

Students read all four stories in a counterbalanced order answering the corresponding set of comprehension questions after completing the reading of each one. The story was available during the answering of the inferential questions in order to control for memory, which would not be an issue for



the automated story comprehension system STAR.

**2.4. Knowledge Elicitation Task**

A knowledge elicitation task was used to prompt readers to activate and record the knowledge they used to comprehend the stories. Each story was presented sentence by sentence and students were asked to write down the knowledge one would need to activate in order to understand the sentence and the story as it unfolded. Story reading, knowledge elicitation, and recording took place on an interactive electronic platform. The word Story along with a number (1, 2, 3, or 4) appeared first on the screen. By pressing the Enter key, students could proceed to read the first sentence along with the story's title. Each subsequent hit of the Enter key would add one more sentence until the final sentence of the story was presented on screen. Proceeding to the next sentence, however, was possible only after students had recorded the knowledge they considered necessary for comprehension by selecting and filling out at least one version of the three basic rule templates available at the bottom of the screen.

In order to facilitate the recording of goal-related, spatiotemporal, and causal knowledge, the available templates represented the basic structure of conditional (henceforth referred to as property) rules (IF ___ THEN ___), and causal rules (___ CAUSES ___) or (___ STOPS ___). In order to facilitate the recording of potential complexity in the knowledge activated, each rule template was available in three versions: simple premise (e.g., IF ___ THEN ___), two-part premise (e.g., IF ___ AND ___ THEN ___), and three-part premise (e.g., IF ___ AND ___ AND ___ THEN ___). Students were allowed to choose and fill out any number and type of the available rule templates for each sentence with a single completed rule template being the minimum requirement.

Two independent raters scored all knowledge protocols in three waves. First, the total number of all non-redundant rules given by each reader to each story was calculated. Any rules that contained only explicit story information were excluded as they did not involve any input from prior knowledge (98 % agreement, Cohen's K = .96, $p < .01$). Subsequently, all rules were examined to determine the extent to which they represented a conceptual or situational association. Rules with premises that reflected specific contexts and situations were classified as situational inferences while rules that reflected generic conceptual relationships that could hold across different contexts and situations were classified as conceptual inferences (see Table 1). Interrater agreement was 84 % (Cohen's K = .68, p < .01), and all differences were resolved through discussion. Finally, all rules were examined as to whether one or more of their premises included an inference or a story element manifested in a previously recorded rule and was, therefore, carried over to the reading of the current sentence. Interrater agreement was 89 % (Cohen's K = .78, p < .01), and all differences were resolved through discussion.

**Table 1**. Examples of conceptual and situational inferences in knowledge protocols

*Conceptual Inferences:*

- IF the doorbell rings THEN someone is at the door.
- Ringing doorbell CAUSES noise.
- IF there is smoke THEN there is fire.
- IF you see something you like THEN you smile.
- IF someone is outside the house AND wants to get in THEN he rings the doorbell.
- The sight of a gun pointing at him STOPS Jim's heartbeat.
- Hunger CAUSES tummy rumbling.

*Situational Inferences:*

- IF the doorbell rings THEN there is a visitor at the door.
- IF someone has a birthday AND people have come to wish him THEN they will ring the doorbell.
- Returning from a fun trip CAUSES a lot of news to share.
- The arrival of the mailman AND the delivery of the new iPad CAUSES enthusiasm.
- IF Lisa is Jenny's roommate AND Lisa helps Sarah THEN Jenny should go help both.
- IF someone rings the doorbell AND wants to surprise someone THEN he does not answer who he is.
- IF Maria got no answer THEN it is either a burglar or a grandmother who does not hear very well.



### 2.5. Procedure

The study was completed in two sessions. Session 1 took place in the students' classrooms and involved the administration of the pre-reading measures (synonym test and story recall task). Session 2 took place about one month later in a computer laboratory. Due to space limitations students participated in two groups. Students were informed that they would be reading a set of stories, one sentence at a time, and answering multiple-choice comprehension questions after each story. The instructions also asked students to use the rule templates at the bottom of the screen to record all the knowledge they deemed necessary for understanding each successive sentence. Before the main experimental session, students practiced reading, knowledge, and answering the multiple-choice inferential questions with a different story that was similar in structure, length, and first sentence with the experimental stories. Subsequently, they proceeded at their own pace reading the experimental stories, recording activated knowledge, and answering the comprehension questions. Session 2 lasted 90 minutes.

## 3. Results

### 3.1. Preliminary Analyses

Accidental keystrokes that required a restart of the interactive electronic platform resulted in 11 students being unable to complete the knowledge recording task, and, therefore, their data were excluded from analyses. Data from the knowledge activation protocols (sample N = 57) were analysed by subject and by sentence across stories (N = 30 sentences). Main dependent variables included measures of amount of knowledge activated (Total and Average Number of Rules), type of association between text input and knowledge reflected in rules (Proportion of Conceptual and Situational Inferences), and continuity (rules including carried-over premises) between current and previous rules and text input (Proportion of Continuous and

Non-continuous Rules). Secondary dependent variables included type and complexity of rules (Proportions of IF-THEN and CAUSES / STOPS rules, and Proportions of Simple and Complex rules). All main and secondary dependent variables were normally distributed (skewness < 1).

One-way ANOVAs indicated that Year (freshman vs. junior) and Department (psychology vs. computer science) had no significant effects on any of the dependent variables (all $p > .05$) and were excluded from further analyses. As expected, Story Comprehension — as indicated by the proportion of correct responses to the inferential questions across stories (M = .81, SD = .11, Mdn = .83) — was relatively high and correlated significantly with Story Recall ($r = .31$, $p = .019$). However, neither of these measures nor Vocabulary knowledge correlated significantly with any of the dependent variables (all $p > .05$), and, therefore, were excluded from subsequent analyses.

**3.2. Amount of Activated Knowledge (Q1)**

Students gave an average of 36.7 (SD = 10.53) non-redundant inference-containing rules across stories (89.3 % of all rules recorded), and an average of 1.22 (SD = 0.36) rules per sentence. There was a higher proportion of property (IF-THEN) than explicit causal rules (M = .81, SD = .14 and M = .19, SD = .14 respectively, paired t (56) = 16.37, $p = .000$), and a higher proportion of Simple than Complex rules (M = .65, SD = .14 and M = .34, SD = .14 respectively, paired t (56) = 8.12, $p = .000$). A series of paired t-tests indicated that number, type, and complexity of rules did not vary significantly across stories (all $p > .05$).

**3.3. Structure of Activated Knowledge (Q2)**

Although students generated more rules containing Conceptual than Situational Inferences (M = .54, SD = .16 and M = .47, SD = .16 respectively) the differences were not significant (paired t (55) = 1.70, $p = .094$). Correlational analyses, performed to examine any associations between kinds of inferences and complexity in rules, indicated that Situational Inferences were positively and significantly correlated with Complex Rules ($r = .31$, $p = .021$), whereas Conceptual Inferences were positively and significantly correlated with Simple Rules ($r = .33$, $p = .000$). In order to examine any Sentence Position effects on the type of inferences generated, multivariate analysis of variance at the item level was performed. Neither the omnibus test (Wilk's $\Lambda = 0.93$, $p = .455$, $\eta2 = .07$) nor the follow-up ANOVA tests indicated any significant effects of Sentence Position on Conceptual Inferences ($F(2, 29) = 1.55$, MSE = 0.03, $p = .231$, $\eta2 = .10$ ) or on Situational Inferences ($F(2, 29) = 1.60$, MSE = 0.03, $p = .221$, $\eta2 = .11$). However, it can be seen from Table 2 that a significantly higher proportion of Conceptual Inferences were generated during the reading of the beginning part of the story (first two sentences) than the middle and the ending parts (Cohen's d = 1.74, Table 2).



**Table 2.** Mean Proportion of Conceptual and Situational Inferences as a Function of Sentence Position across Stories

| Sentence Position | Conceptual | Situational | paired t | p |
|---|---|---|---|---|
| Beginning (n = 8) | .60 (.11) | .40 (.12) | 2.53 | .039 |
| Middle (n = 11) | .46 (.22) | .54 (.22) | - 0.69 | .508 |
| Ending (n = 11) | .51 (.18) | .49 (.18) | 0.15 | .883 |

Note. Standard deviations are shown in parentheses.

**3.4. Function of Activated Knowledge (Q3)**

Overall, students generated a significantly higher proportion of Non-Continuous Rules than Continuous Rules (M = .64, SD = .16 and M = .36, SD = .16 respectively, paired t (56) = -6.38, $p = .000$). In order to examine any effects of Sentence Position on the proportion of Continuous Rules, the first sentence of each story was treated as a missing value and the rest of the sentences were re-classified as occurring either in the middle part story or the ending part of the story. The rationale for this re-classification was based on conceptual and methodological considerations: carried-over elements are not possible during the processing of the first sentence of the story and the inclusion of the data obtained for the first sentence resulted in severe variance heterogeneity across the three sentence positions ($p < .05$). One-way ANOVA for Non-Continuous Rules indicated that the effect of Sentence Position did not reach significance ($F(1, 25) = 3.81$, MSE = 0.02, p = .063, η2 = .14). However, it can be seen from Table 3 that a significantly higher proportion of Non-Continuous Rules were generated only during the reading of the ending parts of the stories (paired t (10) = 3.09, p = .012).

**Table 3.** Mean Proportion of Continuous and Non-Continuous Rules as a Function of Sentence Position across Stories

| Sentence Position | Continuous | Non-Continuous | Cohen's d |
|---|---|---|---|
| Middle (n = 15) | .46 (.14) | .54 (.14) | 0.57 |
| Ending (n = 11) | .35 (.16) | .65 (.16) | 1.88 |
| Cohen's d | 0.73 | 0.73 | |

Note. Standard deviations are shown in parentheses.

In order to examine the possibility of an association between Continuity and Situational Inferences, a systematic sample of 232 rules out of the total 1182 rules was obtained by selecting each fifth rule in the set and excluding rules generated in connection with the first sentence of each story. A Chi-square test for independence (corrected for continuity) indicated significant association between Continuity and Type of Inference in rules, $\chi^2$ (1, *n* = 232) = 29.74, *p* = .000, phi = .37. Specifically, 69 % of the systematically sampled rules were either continuous and situational or non-continuous and conceptual in terms of the carry-overs and the types of inferences contained. Analyses by subject (*N* = 57) confirmed a significant positive correlation between the proportion of Continuous Rules and the proportion of Situational Inference Rules (*r* = .51, *p* = .000).

## 4. Discussion

This study focused on the background knowledge that is activated to support comprehension, and attempted to specify further various aspects of it. Specifically, we sought to examine the amount and the structural and functional characteristics of the knowledge that readers activate for story comprehension. Overall, the findings indicated that readers consciously activate only a limited amount of knowledge that is both conceptually and



situationally connected to input. Conceptual inferences, however, were more prominent than situational inferences only during the processing of the beginning parts of the stories. On the other hand, situational inferences were positively associated with rule complexity, reflecting the activation of larger chunks of knowledge. Notably, the activated knowledge manifested more uniqueness than continuity with previously activated knowledge or text ideas. This, however, was clearly the case only during the processing of the final parts of the stories where unexpected events or situations were introduced. Finally, continuity was associated with situational connections as previously activated ideas that were carried over to current processing were more likely to represent situational than conceptual inferences.

### 4.1. Implications and Limitations

The average number of knowledge-based inferences generated in response to each story sentence ranged from one to two indicating that our readers were rather restrained or selective in the amount of knowledge they recorded in the form of rules. Although this finding in conjunction with the greater number of simple, as opposed to complex, rules in the knowledge protocols agrees with notions of cognitive limitations (Cowan, 1988), it remains rather uninformative regarding online comprehension processes and related theorizing. The knowledge recording task we employed was essentially an inference elicitation task aiming to capture the knowledge deemed relevant for comprehension by a group of readers as opposed to a programmer familiar with the computational needs and specifications of an automated system. Therefore, our methodology did not distinguish between possible construction and integration phases (Kintsch, 1988). Moreover, the possibility of the activated knowledge reflecting integrative or validating processes or both remains open (e.g., Long & Lea, 2005; Singer, 2013).

Although our primary aim was to harvest the knowledge that readers use to comprehend stories with the intent to represent it computationally, it must be noted that the nature of the knowledge recording task we employed may have imposed artificial limits to the amount of the knowledge obtained. As implicated above, the task could not capture any knowledge that was automatically activated in the course of comprehension. More importantly, however, the possibility of knowledge consciously activated but not reported remains open. This omission may have been due either because readers deemed the activated knowledge too trivial or too cumbersome to report given the constraints of the templates that were available to them. The knowledge templates expressed a form of implication (be it causal or not) from a conjunctive premise to a single conclusion. Knowledge that would require more complex or non-conjunctive premises or that would require a more involved conclusion (e.g., a disjunction of two possibilities) would likely not be reported. Moreover, we cannot preclude the possibility that simply the order of presentation of rule templates in the knowledge recording task hindered the selection of more complex rules and causal rules which always appeared, respectively, after simple rules and property rules. Nevertheless, also unreportable would be knowledge of a non-implicational form, such as the similarity of a story part with another part of that story or of a different story, the simple association of a concept with another concept, any inductive knowledge that might result from repetitive elements in a story, or knowledge about the expectations of the participants when reading a story (and more so, in the context of a psychological study). Without diminishing the importance of such type of unreportable (given the templates in our study) knowledge, excluding the reporting of such activated knowledge was part of the study's aims: the study sought to investigate the characteristics of the activated knowledge that is, precisely, of a form (as given in the templates) that would be directly usable by the computational model that guided the design of the study. Future work that employs online but informative measures, such as probe recognition, is needed to establish the sufficiency or not of this type / structural form of knowledge for certain story comprehension tasks.

The findings regarding the structural characteristics of the activated knowledge can be viewed in two ways. First, the overall greater reliance on conceptual than situational associations between story input and prior knowledge point in the direction of relatively fine-grained knowledge being activated in response to story input. This is in line with the general assumptions and predictions of comprehension models that postulate the activation of related concepts via spreading activation or resonance mechanisms (Kintsch 1988; Myers & O'Brien, 1998). Second, the lower but still substantial amount of situation-specific inferences generated at the very



beginning of the stories is also in line with a scenario-mapping hypothesis (Sanford & Garrod, 1998) and evidence suggesting that knowledge-based inferences are generated as soon as the linguistic context makes available the necessary information (Cozijn, Noordman, & Vonk, 2011; Dery & Koenig, 2015; Matsuki, Chow, Hare, Elman, Scheepers, & McRae, 2011). Moreover, the association between situational inferences and rule complexity (rules with multiple premises) is in line with the notion that situational inferences reflect the activation of larger chunks of functionally interrelated knowledge. We must note, however, that the restricted sample size (in subject and item analyses) may have prevented differences from reaching significance, while the inference elicitation task may have inflated the amount of conceptual or situational inferencing observed.

Nevertheless, the general pattern of findings regarding the effects of the story part being processed on the kinds of knowledge activated is clearer. The reliance on general conceptual associations was clear and significant during the processing of the beginning parts of a story only. This result in conjunction with the notable but non-significant increase in conceptual associations during the processing of the final part of the story is consistent with the predictions of both the structure-building (Gernsbacher & Roberston, 1992) and event-indexing models of comprehension (Zwaan, et al., 1995). Therefore, the increase in situational inferences during the processing of the middle story parts may reflect the incremental building of a situation model that can plausibly encompass all the information derived from the input (e.g., Kurby & Zacks, 2012). Nevertheless, the fact that situational inferences were also elicited during the processing of the first sentence in the stories indicates the potential availability of situation-specific knowledge whose full activation may be a function of a complex interaction between person, text, and task.

The results concerning the coherence function of the activated knowledge are consistent with a theoretically predicted pattern of increased discontinuity during the processing of the ending parts of the stories (e.g., Zwaan et al., 1995). The events described in these ending parts occur in the same time and location as the previously read ones and they may or may not introduce changes in the goals of the protagonist or additional entities (see Appendix). However, they consistently represent breaks in the causal chain of events in the sense that they cannot reflect the direct effects of the actions or the situations described in the previous story parts. As such, they violate any expectations that readers may have formed so far in their story reading necessitating the (at least partial) deactivation of knowledge that gave rise to those expectations. Moreover, causal breaks represent event boundaries that are more consistently associated with increased processing time than temporal, spatial or entity shifts in the story line (Pettijohn & Radvansky, 2016). This finding underscores the significance of causal coherence building in the comprehension of stories and of causal breaks as triggers for situation model updating or revision. Although the association of discontinuity with causal breaks documented with offline measures in this study is in line with the online results of Pettijohn and Radvansky (2016), they also raise important questions. First, the possibility of discontinuity being associated with other kinds of changes in the story line (e.g., spatiotemporal shifts) remains open. Second, in our study, as in Pettijohn and Radvansky (2016), a causal break is conceptualized to occur when subsequent events are incompatible with the predicted or predictable effects of previous events. Unpredictable events in a story context are likely to coincide with or entail changes in protagonist goals. For example, in our second story the goal of the protagonist is bound to change from enjoying his new gadget to protecting himself in a dangerous situation (see Appendix). Disentangling the potential effects of causal and goal changes on processing and knowledge activation appears to be more feasible in cases when a causal break involves effects of subsequent events that are incompatible with those of previous events, not the events themselves. In these cases, causal and goal changes may or may not coincide. For example in our first story, the main goal of the protagonist was frustrated temporarily but not changed by the appearance of the neighbour (see Appendix). Examining the extent to which discontinuity in the activated knowledge is associated with other changes in the story line requires further research that introduces more systematic control of the different types of changes in the line of stories (Zwaan et al., 1995).

It is noteworthy, however, that the overall continuity manifested in the knowledge activated during the processing of successive sentences was low. In general, our readers were more likely to activate new knowledge



in response to each sentence than to carry over previously activated or encountered ideas. We must note, however, that the combination of simple short stories that posed no challenge for maintaining coherence and a task that required readers to activate as much knowledge as they considered necessary for their comprehension may have conspired to reduce the need for or the manifestation of continuity. Nevertheless, continuity was associated with the activation of situation-specific knowledge. Situational inferences were more likely than conceptual inferences to be carried over for the processing of a following story segment. As a result, overarching themes or scenarios afforded by the stories (e.g., visiting, parcel delivery, or Christmas carolling) tended to be repeated across rules while generic conceptual associations (e.g., person ringing doorbell) tended to be left behind.

The association between situational inferences and continuity can be viewed to indicate that knowledge representing multiple and complex functional relationships that ground story information in world knowledge and experience contribute to the building of representational coherence to a greater extent than generalizable conceptual associations. However, considering that activated situation-specific knowledge may provide a framework for explaining and assessing the plausibility of information encountered in a story (Connell & Keane, 2006; Graesser et al., 1994; Singer, 2013), this finding may also reflect readers' unwillingness to abandon a situational framework that keeps on serving their comprehension well. This possibility carries implications for comprehension models predicting fluctuations in the activation strength of concepts in a network representation as a result of successive input (e.g., Kintsch, 1988; Tzeng, et al., 2005). Specifically, the activation strength of a node may also be a function of its content (what it represents) in addition to what the input dictates. Nevertheless, further research employing a more refined and targeted methodology is needed to examine the potential connections between knowledge activation and representational coherence building.

To summarize, the study indicated that although the amount of knowledge readers activate is limited, there is substantial variability regarding its structural characteristics and coherence function. Our readers demonstrated flexibility in activating both general conceptual and situation-specific knowledge depending on the part of the story they were processing. Moreover, they were keener to activate unique knowledge than to carry over previously activated ideas. Nevertheless, situation-specific knowledge was more likely than fine-grained conceptual knowledge to remain active for the processing of subsequent story segments.

### 4.2. Implications for the Computational Representation of Knowledge

Although the participants were instructed, and did not freely choose, to represent their activated knowledge in the form of property and causal rules, the variability of the activated knowledge in terms of its structural characteristics and coherence function that was successfully represented, and the prominence of non-redundant inference-containing rules in the activated knowledge, offers some evidence on the psychological validity of the uniform representation of knowledge in the form of rules that can then be computationally reasoned with using standard or specialized computational rule-based inference mechanisms.

The limited amount of activated knowledge can be taken as an indication that the process of elaboration in humans is rather simple and ultra-efficient, which might seem to clash with the possibly natural but still nuanced semantics of certain computational story comprehension systems or approaches (e.g., Diakidoy et al., 2014; Frank et al., 2003). However, the nuance in computational story comprehension can be attributed to the need to address another task, for which the current study offers no explicit feedback: the task of selecting which knowledge to activate among the potentially significant body of knowledge that might be available. A computational selection mechanism needs to deal not only with the identification of knowledge that is relevant for a given story, but also with the identification of the subset of that knowledge that is coherent (and self-consistent) and respects any resource constraints that the automated system may have. Critically, the selection mechanism needs to be able to revise its selection in the presence of story surprises. We posit that if this extra task that computational systems need to undertake is somehow isolated from the elaboration process with the activated knowledge, then the latter process can indeed be mechanized in a way that agrees with the study findings, by employing only a small fraction of all available knowledge, and by



applying a simple efficient inferencing mechanism.

The finding on the prominence of property rules over causal rules, suggests that the automated acquisition of a significant part of the knowledge needed to comprehend stories could be carried out using symbolic machine learning over textual corpora to extract associations between frequently co-occurring words or propositions (Michael, 2009; Michael & Valiant, 2008), in line with the distributional hypothesis (Firth, 1957). The additional finding on the prominence of simple rules offers further support that a learning-based acquisition process is appropriate: rule simplicity implies a smaller hypothesis space for learning, and therefore a more efficient acquisition process. Although an analogous approach could be potentially useful for the acquisition of causal rules, we expect that script-related textual corpora, and not general textual corpora, would be better as training material for the acquisition of causal knowledge. This appeal to script-related textual corpora would be further supported by a potential finding suggesting a correlation of causal rules with situational rules, since script-related textual corpora are more likely to include information on the temporal evolution of a situation. Although our study hinted towards such a possible finding, the results were not statistically significant, and were not reported; a larger scale study would be needed to support such a finding and the implication on the knowledge acquisition process. It also remains open whether knowledge important for story comprehension that was potentially activated but not reported by human readers in the present study may be analogously acquired from appropriate textual corpora. The current study offers no explicit evidence on the nature of this unreported but potentially activated knowledge, and a new study quantifying its characteristics would be a necessary first step.

In terms of the coherence function of the activated knowledge, the prevalence of non-continuous rules, along with the fact that the majority of the activated knowledge was in the form of property and not causal rules, may suggest a limited or careful activation of temporal reasoning during story comprehension. This, again, might seem to be in contrast to the reasoning semantics of existing computational systems in the common sense reasoning literature that attempt to carefully account for various temporal phenomena (see, e.g., Mueller, 2014). As we have argued earlier, this discrepancy could be an artefact of the extra knowledge-selection process that these computational systems incorporate. Although the study does not explicitly investigate this selection process in human readers, it does suggest a potential heuristic that computational story comprehensions systems could adopt for more efficient temporal reasoning: When a surprise is encountered, stop carrying over situational inferences and retain only conceptual ones. Combined with a mechanism to automatically distinguish situational from conceptual inferences (which in the present study was carried out by trained psychologists) this heuristic could improve the reasoning efficiency of the systems at the expense of occasionally drawing wrong inferences in complex temporal reasoning settings. The correlation of situational rules with complex rules already offers one hint towards developing such heuristic that can be empirically tested: the inferences coming from more complex rules are less likely to carry over once a surprise is encountered. If additional studies establish a correlation between situational rules and certain other automatically quantifiable characteristics of rules (such as their causal nature), these characteristics could also offer the basis for the development of additional heuristics.

To implement that above heuristic, one needs to automatically identify story surprises, which could be naturally associated with discontinuities in the story. Such discontinuities, however, do not always suggest a situational change (e.g., the same situation could persist despite a spatiotemporal shift), or might require the stopping of persistence of some aspects of a situation, and the retention of the other ones (e.g., change of a protagonist's position but not the protagonist's goal). We posit that meta-knowledge on the effects that story discontinuities have on the persistence of inferences and story elements across the story timeline can be represented in the same form as object-level knowledge.

### 4.3. Conclusion

The study focused on the knowledge activated by readers during story comprehension with the intent to inform efforts for computational knowledge representation and story comprehension. Therefore, it explored issues that reflect important considerations in a computational context, namely, how much knowledge and of what form would support automated comprehension in a way that allows for coherence building and revision capability. The findings point in the direction of representation in the form of association rules that can capture a variety of relations critical for story comprehension. This knowledge representation and reasoning approach acknowledges the argumentative nature of human reasoning and accommodates computationally the sparseness of the activated knowledge during comprehension. More importantly, however, the documented variability regarding the grain-size, situational specificity, and persistence of the activated knowledge as the story unfolds represent aspects that a computational system ought to take into account as well as the criteria against which it can be validated. A follow-up study could seek to characterize the parts of the knowledge activated by humans in the present study that are actionable (i.e., are actually invoked) in the process of answering questions in a computational story comprehension system like STAR, and also characterize any extra knowledge that might be required.



## Acknowledgments

We thank Christos Rodosthenous for the design of the interactive electronic platform and Sylvia Savvidou and Melina Ioannou for their help with data collection and scoring.
Correspondence concerning this article should be addressed to Irene-Anna N. Diakidoy (ORCID: 000-0003-3650-8108), Department of Psychology, University of Cyprus, 1 University Avenue, Aglantzia, 2109 Cyprus. E-mail: eddiak@ucy.ac.cy

## References

Anderson, R. J. 1993. Problem solving and learning. American Psychologist, 48, 35-44.

Bench-Capon, T.J.M., & Dunne, P.E. 2007. Argumentation in Artificial Intelligence. Artificial Intelligence, 171(10-15), 619-641, DOI: 10.1016/j.artint.2007.05.001

Budiu, R., & Anderson, J. R. 2004. Interpretation-based processing: A unified theory of semantic sentence comprehension. Cognitive Science, 28, 1-44, DOI:

10.1016/j.cogsci.2003.10.001

Connell, L., & Keane, M. T. 2006. A model of plausibility. Cognitive Science, 30, 95-120.

Cook, A. E., Halleran, J. G., & O'Brien, E. J. 1998. What is readily available during reading? A memory-based view of text processing. Discourse Processes, 26, 109-129, DOI: 10.1080/01638539809545041

Cowan, N. 1988. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. Psychological Bulletin, 104, 163-191, DOI: 10.1037/0033-2909.104.2.163

Cozijn, R., Noordman, L., & Vonk, W. 2011. Propositional integration and world-knowledge inference: Processes in understanding 'because' sentences. Discourse Processes, 48, 475-500, DOI: 10.1080/0163853X.2011.594421

Dery, J. E., & Koenig, J.-P. 2015. A narrative-expectation-based approach to temporal update in discourse comprehension. Discourse Processes, 52, 559-584, DOI: 10.1080/0163853X.2014.966293

Diakidoy, I. N., Kakas, A., Michael, L., & Miller, R. 2014. Story comprehension through argumentation. In S. Parsons, N. Oren, C. Reed, and F. Cerutti (Eds.), Frontiers in Artificial Intelligence and Applications, vol. 266 (pp. 31-42). IOS Press, DOI: 10.3233/978-1-61499-436-7-31

Firth, J.R. 1957. A synopsis of linguistic theory 1930-1955. Studies in Linguistic Analysis. Oxford: Philological Society: 1–32. Reprinted in F.R. Palmer, ed. 1968. Selected Papers of J.R. Firth 1952-1959. London: Longman.

Frank, S. L., Koppen, M., Noordman, L. G., & Vonk, W. 2003. Modeling knowledge-based inferences in story comprehension. Cognitive Science, 27, 875-910, DOI: 10.1016/j.cogsci.2003.07.002

Frank, S. L., Koppen, M., Noordman, L. G., & Vonk, W. 2008. World knowledge in computational models of discourse comprehension. Discourse Processes, 45, 429-463, DOI: 10.1080/01638530802069926

Gernsbacher, M. A., & Robertson, R. R. W. 1992. Knowledge activation versus sentence mapping when representing fictional characters' emotional states. Language and Cognitive Processes, 7, 353-371.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. 1997. Discourse comprehension (pp. 163-189) . In J. T. Spence, J. M. Darley & D. J. Foss (Eds.), Annual review of psychology (Vol. 48). Palo Alto, CA: Annual Reviews.

Graesser, A. C., Singer, M., & Trabasso, T. 1994. Constructing inferences during narrative text comprehension. Psychological Review, 101, 371-395.

Kintsch, W. 1988. The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95, 163-182.

Kintsch, W. 1998. The representation of knowledge in minds and machines. International Journal of Psychology, 33, 411-420.



Kintsch, W., & Mangalath, P. 2011. The construction of meaning. Topics in Cognitive Science, 3, 346-370, DOI: 10.1111/j.1756-8765.2010.01107.x

Kowalski, R., & Sergot, M. 1986. A logic-based calculus of events. New Generation Computing, 4(1), 67-95, DOI: 10.1007/BF03037383

Kurby, C. A., & Zacks, J. M. 2012. Starting from scratch and building brick by brick in comprehension. Memory and Cognition, 40, 812-826, DOI: 103758/s13421-011-0179-8

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.

Long, D. L., & Lea, R. B. 2005. Have we been searching for meaning in all the wrong places: Defining the "search after meaning" principle in comprehension. Discourse Processes, 39, 279-298.

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. 2011. Even-based plausibilityimmediately influences on-line language comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37, 913-934, DOI: 10.1037/a0022964

McCarthy, J. 1963. Situations, actions and causal laws. Technical report, Stanford University. Reprinted in Semantic Information Process (M. Minsky ed.), MIT Press, Cambridge, Mass. 1968, pp. 410-417.

McNamara, D. S., & Magliano, J. 2009. Toward a comprehensive model of comprehension. In B. Ross (Ed.), The psychology of learning and motivation (Vol. 51, pp. 297-384). Burlington: Academic Press.

Mercier, H. & Sperber, D. 2011. Why do humans reason? Arguments for an argumentative theory. Behavioral and Brain Sciences, 34, 57-111, DOI: 10.1017/S0140525X10000968

Michael, L. 2009. Reading between the lines. Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI), pages 1525-1530, from the website: http://www.aaai.org/ocs/index.php/IJCAI/IJCAI-09/paper/view/512

Michael, L. & Valiant, L. G. 2008. A first experimental demonstration of massive knowledge infusion. Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR), pages 378-389, from the website: http://www.aaai.org/Library/KR/2008/kr08-037.php

Mueller, E. T. 2014. Commonsense reasoning: An Event Calculus based approach. 2nd edition. San Francisco, US: Morgan Kaufmann

Myers, J. L., & O'Brien, E. J. 1998. Accessing the discourse representation during reading. Discourse Processes, 26, 131-157, DOI: 10. 1080/01638539809545042

Norman, D. A. 1982. Learning and memory. New York, NY: Freeman.

Pettijohn, K. A., & Radvansky, G. A. 2016. Narrative event boundaries, reading times, and expectation. Memory and Cognition, 44, 1064-1075, DOI: 10.3758/

s13421-016-0619-6

Radvansky, G. A., & Copeland, D. E. 2010. Reading times and the detection of event shift processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36, 210-216, DOI: 10.1037/a0017258

Rahwan, I., & Simari, G. 2009. Argumentation in Artificial Intelligence, Berlin: Springer-Verlag.

Sanford, A. J., & Garrod, S. C. 1998. The role of scenario mapping in text comprehension. Discourse Processes, 26, 159-190.

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

Schunn, C. D., & Klahr, D. 1999. Production systems. In W. Bechtel and G. Graham (Eds.), A companion to cognitive science (pp. 542-551). Oxford, UK: Blackwell.

Singer, M. 2013. Validation in reading comprehension. Current Directions in Psychological Science, 22, 361-366, DOI: 10.1177/0963721413495236

Singer, M., & Kintsch, W. 2001. Text retrieval: A theoretical exploration. Discourse Processes, 31, 27-59, DOI: 10.1207/S15326950dp3101_2

Smith, E., & Hancox, P. 2001. Representation, coherence and inference. Artificial Intelligence Review, 15, 295-323.

Suh, S., & Trabasso, T. 1993. Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. Journal of Memory and Language, 32, 279-300.

Toolan, M. 2006. Narrative: Linguistic and structural theories. In K. Brown (Ed.), Encyclopedia of language and linguistics (2nd ed., pp. 459-473). Amsterdam, NL: Elsevier.

**Trabasso**, T., & **van den Broek**, **P**. **1985**. Causal thinking and the representation of narrative events. Journal ofMemory and Language, 24, 612–630.

Trabasso, T., & Wiley, J. 2005. Goal plans of action and inferences during comprehension of narratives. Discourse Processes, 39, 129-164, DOI: 10.1207/S15326950dp3101_2

Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C. 2005. The computational implementation of the landscape model: Modeling inferential processes and memory representations of text comprehension. Behavior Research Methods, 37, 277-286.

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. 2001. The effects of readers' goals on inference generation and memory for texts. Memory & Cognition, 29, 1081-1087.

Zwaan, R. A. 1996. Processing narrative time shifts. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 1196-1207.

Zwaan, R. A. 2016. Situation models, mental simulation, and abstract concepts in discourse comprehension. Psychonomic Bulletin & Review, 23, 1028-1034, DOI: 10.3758/s13423-015-0864-x

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(5), 292-297.

## APPENDIX

Experimental stories and examples of inferential questions

**Story 1: A case of absentmindedness**

The doorbell rang. Helen smiled. She was expecting the doorbell to ring many times tonight. She took her little basket with the chocolates and went to the door. She enjoyed very much listening to Christmas carols at her door. She opened the door and saw Eve, her neighbour, standing outside and looking embarrassed. "Not again!" Helen thought. Feeling a bit angry, Helen went to the kitchen to bring Eve's house keys that she kept for such occasions.

Example Question: Why was Helen feeling a bit angry? Because

a) she did not want to give Eve a chocolate.
b) Eve was bothering her often.
c) there were no children at the door.

**Story 2: Nasty surprise**

The doorbell rang. Jim jumped up spilling his juice on the table. He did not expect any visitors that early in the morning. But finally, the time had arrived. Soon he would be holding his new iPad and exploring its capabilities. He rushed to open the door and found himself staring at a gun pointing at him. A frightened stranger was signing him to be quiet while pushing him inside the apartment.

Example Question: Whom did Jim expect to see at the door?

a) The mailman.
b) The electrician.
c) The neighbour.

**Story 3: Night snack**

The doorbell rang. Jenny opened her eyes, stretched, and yawned. She walked lazily out of the bedroom when the doorbell rang again. She passed through the door and went to the kitchen. Her tummy was rumbling, and she was looking for something to eat. When she got out of the kitchen a bit later, Lisa was standing at the door and helping Sarah bring in her luggage.



After they talked for a while, they said "goodnight" and went to bed. Jenny, having filled her tummy, went happily to continue her sleep.

Example Question: Who is Jenny?

a) Lisa's roommate.
b) Lisa's pet.
c) Sarah's cabdriver.

**Story 4: Home emergency**

The doorbell rang. Maria asked who it was, but got no answer. She was alone at home and was afraid to open the door. She walked toward the door and stopped. She could smell smoke from the outside. Maria was trapped in her apartment. Now she was definitely afraid to open the door.

Example Question: Why was Maria trapped in her apartment? Because

a) she did not have the keys.
b) there was no emergency exit.
c) there was a fire outside her apartment.

Note. The stories were translated from the Greek language. The middle part of each story is shown in italics here, and only for presentation purposes.