Pickup of Causal Language and Inference

During and After Reading Scientific Text

Accepted version of Cromley, J. G., Ma, S., Van Boekel, M., & Dane, N. (2020). Pickup of causal language and inference during and after reading illustrated text. Reading Psychology, 41(3), 157-182. https://doi.org/10.1080/02702711.2020.1768974

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A160335 to the University of Illinois, Urbana-Champaign. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract

When reading scientific text, readers must draw inferences when the author does not make relations explicit; readers also need to pick up on causal relations that the author *does* make explicit. We collected think-aloud protocols from 86 undergraduate biology students reading 7 brief, illustrated passages about the immune system. After reading, participants typed a free recall. Both think-aloud and recall were transcribed and coded for two types of causal language—specific action verbs and explicit causal terms (forms of the word "cause")—as well as inferences. Inferences in recalls were significantly related to action verbs picked up and during reading, as evidenced by the think-alouds. Action verbs in recalls were significantly facilitated by inferences and pick up of action verbs articulated by the participants while reading, as evidenced by the think-alouds. Explicit causal terms in recalls were significantly facilitated by pick up of action verbs while reading. Results suggest a very important role for pick-up of specific action verbs while reading, and risks associated with learners substituting vague verbs of class membership ('is a,' 'has a') for highly specific action verbs ('activates,' 'binds') when learning biology. One possible implication is that prompting students to pick up on specific action verbs while reading biology texts—in addition to making inferences during reading—could prompt more benefit from reading these types of texts.

Pickup of Causal Language and Inference During and After Reading Scientific Text

Inference making has been an important focus in much of reading comprehension research, as well as being a focus of major theories of comprehension (e.g., the Construction-Integration Model, the Landscape Model; O'Brien, Cook, Lorch, & Lorch, 2015). When reading scientific text, readers must draw inferences to facilitate understanding when the author does not make causal relations explicit (Otero, Leon, & Graesser, 2002); readers also need to pick up on causal relations that the author *does* make explicit to aid further recall of information. Previous research has shown that making inferences during reading is beneficial, as it is associated with forming a better mental model of the depicted situation (Butcher, 2006; Kintsch, 1998; McNamara, 2004). However, not much is known about to what extent learners pick up on the language specifically related to causality—both explicit causal terms (e.g., forms of "because") and action verbs—from text, and to what extent this pick up is reflected in mental model construction. Therefore, the present study investigates the association between pick up of causal language while reading and the subsequent recall of the information contained in the text. In order to accomplish this goal, participants were asked to think-aloud while reading a series of short scientific texts and upon completion were given a surprise memory task where they engaged in free recalls for the texts.

Causal language refers to words or phrases that one uses to express causality, i.e., the change of an event or an object leads to the change of another event or object (Koslowski & Masnick, 2010). Explicitly stated causal language (e.g., histamine *opens holes in* blood vessels) are common in scientific text, and comprise action verbs or causal terms (e.g., forms of "cause" or "because," if…then, as a result [of], due to), as opposed to state-of-being verbs (e.g., is next to, is made up of). The presence of action verbs represents instances where the author has explicitly provided information about how different structures or substances act on each other to make a change in other substances/structures or in the whole system (in the example above, histamine acts on blood vessels by opening holes in them). The presence of causal terms indicates underlying causal relations between actions of different substances or structures. When the author makes such relations explicit, the reader does not need to make an inference, but at the same time the reader needs to attend to the specific nature of action and causality.

Structures (elements) and functions (relations) in biology comprehension

In the present study, we focus on readers' comprehension of biological texts. A major theme in biology as a scientific discipline is the relationship between biological structure (things, represented by nouns) and biological function (relations, represented by verbs). That is, particular biological elements and their inter-relationships are directly related to the functions of those parts. For example, the constant and variable regions of heavy and light polypeptide chains *attach* to receptors on immune system cells, *recognize* and *bind* to unique antigens. Undergraduate biology students are known to have difficulty with learning structure-function connections, in part because high school biology is often taught with an emphasis on structures (College Board, 2015). Indeed, the passages we gave participants to read in this study did vary in their relative emphasis on structures (state-of-being verbs) and functions (action verbs). Thus, we would not be surprised if our undergraduate biology student participants paid more attention to structures/substances (nouns) than to functions (action verbs) while reading.

There is little research that separates students' learning of structures from their learning of functions, or the extent to which they learn structure-function relations. Early research by Lambiotte and Dansereau (1992) presented all participants with brief videotaped biology lectures accompanied with a randomly assigned lecture aid: a concept map (which explicitly included structures and functions and labeled causal relations between them), a textual outline (which also explicitly included structures and functions but did not label causal relations), or a list of key terms (which did not separate structures and functions, and had almost no action verbs at all). Although post-video recall did not separate out structures from functions, the 'noun-focused' list condition led to the highest proportion of isolated, disconnected facts in recall. Focusing on structures in instruction was associated with missing the structure-function relation when learning these biology topics. A focus on isolated, disconnected facts when reading should lead to poor comprehension and poor learning from text. By contrast, noticing or making links between facts—such as structure-function relations—should lead to better comprehension and better learning from text.

To foreground our findings, we conclude that participants form a very poor textbase (i.e., a semantic understanding of the text; Kintsch, 1998) when they over-emphasize structures and under-emphasize functions. Indeed, evidence from the think-aloud protocols we collected suggest that some participants form mental models where various facts are linked via extremely simple connections such as 'has' or 'includes' (e.g., the immune system includes antigens and antibodies, T cells have receptors). The overly-general links mean there is an impoverished textbase which misses most information about how these various facts are related (e.g., antibodies signal the presence of antigens in the body).

Construction-integration, representations of text, and specificity

Kintsch's (1998) Construction-Integration model provides us with a way to rank order the types of verbs that comprise a participant's mental model, from overly-general ('includes') to highly specific ('signals'). In Kintsch's (1998) Construction-Integration (CI) Model, readers form different representations of text (each of which can range from lower- to higher-quality) depending on the extent to which they incorporate their own prior knowledge: 1) a surface form or verbatim text model, similar to a 'photographic memory' of the text, 2) a textbase or gist model of the text, which is a summary of what was read, but without adding any information from the reader's prior knowledge, and 3) a situation model, in which information from the text is incorporated with information from the reader's prior knowledge. The situation model is posited to be a higher level of comprehension than the textbase (Graesser & Britton, 1996; Royer, Greene, & Sinatra, 1987). From a practical standpoint, these comprehension quality differences arise because more inferences are generated when forming a situation model than a textbase representation of the text. Further, the during-reading processes (e.g., such as bridging inferences and elaborative inferences) and the reading strategies (e.g., summarizing, self-questioning, making a drawing) required to form a situation model are far more sophisticated than those required for forming a textbase representation of the text.

A critical part of the CI model is the proposition, and as noted by Kintsch (1998), "verb frames…[are] building blocks for propositions" (p. 55). This implies that the level of specificity in the verbs that readers pick up on ('includes,' 'activates') have important implications for the quality of the proposition in memory. For example, a biology student who takes notes that read "histamine, blood vessels" implicitly is using a very vague verb phrase to connect these, along the lines of 'has to do with' ('histamine has something to do with blood vessels'). The student has not picked up on the specific verb in the causal relation "opens holes in" provided by the author, and instead has defaulted to a much more vague 'has a relation to' schema.

When a learner uses vague verbs of general relation and foregoes the use of action verbs when representing the text they are reading, the learner's impoverished representation of the text will result in a vague, low-quality textbase representation of the content. Such an impoverished textbase is not optimal for comprehension and would be expected to interfere with the inference-making required to form a sophisticated situation model. By contrast, when learners do pick up specific action verbs and explicit causal terms during reading, the learner is likely to form a detailed, high-quality textbase that should be optimal for comprehension, supporting future inferences.

Importance of causal language pickup

There is a small body of literature focused specifically on reader pick up (e.g., in a verbal or written paraphrase, notes, or summary) of causal language from presented text or presented concept maps during reading, which is then related to post-reading measures (Cho & Jonassen, 2012; Leon & Escudero, 2015; McMaster et al., 2015; Naumann, Wechsung, & Krems, 2009; Sanchez & Wiley, 2014; Wiley & Voss, 1999; Wiley et al., 2009). In a typical research design, participants are asked to learn from a scientific text, the text is removed, and they are asked to write about the scientific topic. Pick up of causal language from the text(s) is considered a sign of more coherence of the reader's own mental representation (Leon & Escudero, 2015). High school participants who pick up more causal language when asked to self-explain a biology phenomenon score higher on inference questions at posttest (Cho & Jonassen, 2012).

Different presentation modalities are associated with different pick up of provided causal concepts (non-illustrated texts show the same pick up as static diagrams, which show less pick up than animations) for undergraduate non-science majors learning about volcanoes (Sanchez & Wiley, 2014). Undergraduate participants pick up more causal language from text as manifested in written post-reading summaries, compared to high school participants, and also show higher comprehension (Leon & Escudero, 2015). Undergraduates asked to write an argument summarizing multiple documents produced more causal language and scored higher on a posttest inference task than those given explanation, summary, or narrative tasks (Naumann, Wechsung, & Krems, 2009; Wiley & Voss, 1999; Wiley et al., 2009). Struggling 4th grade readers provided with causal prompts recalled more than those given general prompts on near transfer reading tasks, but showed no differences on causal language at posttest (McMaster et al., 2015). These studies together provide evidence that not picking up on the author's causal language is in fact associated with less causal language in recalls, fewer inferences, and poorer comprehension.

This link between causal language and inference can be illustrated with an example from Wiley et al (2009): a participant who has picked up that 'ocean magma *fills* chambers rapidly' and 'chambers *vent* when they are full' can make the inference that 'ocean magma is vented from chambers.' By contrast, a participant who notes 'ocean, magma, chambers' ('these three things are parts of volcanoes') is unlikely to make such an inference.

Roles of verbs, nouns, and causal terms during reading

The roles of action verbs and explicit causal terms are intertwined in discourse. Without action verbs one cannot make a causal statement such as 'NF-kB is able to *enter* [action verb] the nucleus because [explicit causal term] the protein kinase cascade *changes* [action verb] the shape of the NF-kB molecule'. Notwithstanding, readers might pick up only on action verbs ('protein kinase cascade *changes* the shape of the NF-kB, NF-kB *enters* the nucleus) without the explicit causal terms; yet the reverse cannot be true, i.e., there is no way to pick up on explicitly causal terms without also using action verbs ('something happens *because* something else happens'). We therefore argue that picking up on action verbs but not on explicit causal terms would result in a medium-quality textbase, whereas picking up on action verbs and explicit causal terms would result in a higher-quality textbase.

By contrast, notes, paraphrases, or summaries that miss the action verbs tend to consist almost entirely of nouns (e.g., 'protein kinase cascade, NF-kB, nucleus'). Presumably, the reader has formed a textbase with vague verbs of general relation along the lines of 'there is a protein kinase cascade and NF-kB and the nucleus are involved'. Indeed, when students are asked to show their understanding of texts such as these (e.g., on an Advanced Placement biology exam), the 'nouns-only' answer would be marked as inferior to the 'nouns-plus-causal-verbs' answer, but the latter would be marked as inferior to one with nouns, verbs, and inferences that go beyond the stated information (College Board, 2018).

In summary, while inference-making is the goal of reading, and represents high-level cognitive activity, the lower-level tasks of picking up on action verbs and explicitly causal terms are not trivial tasks. Doing so is likely important for making later inferences. Picking up on action verbs and causal terms is particularly relevant for students studying biology because of the field's need to rely on structure-function representations to describe key concepts.

We hypothesize that readers who recognize the importance of causal language during reading—e.g., by including more causal words in a paraphrase—are building a more specific, higher quality textbase (structures/substances AND causal relations vs. only structures/substances) than those who pick up only on structures and substances. To test this hypothesis we asked participants to think aloud as they read seven short pages of biology text on immunology. Each page had an accompanying figure. Once completed, participants were given a surprise memory task and asked to recall everything that they could remember from the texts they had just read. We then coded for action verbs and explicit causal terms separately, as well as inferences made both during reading and at post-reading recall. In this way, we add to the small literature on causal language in comprehension, with a sample of biology majors (higher ecological validity), separating out two different types of causal language (action verbs and explicit causal terms), and measuring these together with inferences.

Research Goals, Hypotheses, and Analysis Plan

We use three regressions to test these relations in this study. In the first regression, we test whether pick up of explicit causal terms and action verbs during reading, and inferences made during reading are associated with explicit causal terms at recall. We expect that the more explicit causal terms readers verbalize while learning, the more explicit causal terms they will verbalize at recall. This expectation is based on the reasoning that readers may have difficulty recalling explicit causal terms that they simply read, but never picked up on while reading. We expect that the more action verbs readers pick up on while reading, the more explicit causal language they could use at recall, since deploying the explicit causal language requires the presence of action verbs. We expect that the more inferences readers make during reading, the more causal language they could use at recall, since a number of these inferences are causal inferences (see the example from Wiley et al., 2009 above).

In the second regression, we test whether pick up of explicit causal terms and action verbs during reading and inferences made during reading are associated with action verbs at recall. We argue that pick up of explicit causal terms made during reading is likely to be associated with the presence of action verbs at recall because these linguistic features are intertwined in discourse (Barriere, 2001). We expect that pick up of action verbs during reading is likely to be associated with generation of action verbs at post-reading recall because readers may have difficulty recalling action verbs that they simply read, but never picked up on while reading. We argue that inferences made during reading may be associated with action verbs at recall because the inference is itself based on the action verbs; possibly, the inference could be forgotten but the action verbs involved in making it could be remembered.

In the third regression, we test whether pick up of explicit causal terms and action verbs during reading and inferences made during reading are associated with inferences at recall. We expect that the more explicit causal terms participants verbalize while reading, the more inferences they will make at recall. Notably, this may be dependent on pick up of action verbs as well, but inferences by their nature also rely on explicit causal terms, and we reasoned that participants who pick up on explicit causal terms while reading will more easily be able at recall to draw the inferences that involve those causal terms. This hypothesis is supported by four studies reviewed above (Cho & Jonassen, 2012; Author et al., 2010; Leon & Escudero, 2015; McMaster et al., 2015). Inferences by their nature rely on action verbs, and we reasoned that participants who pick up on action verbs while reading will have the pre-requisite information to make inferences (e.g., histamine *opens holes in* blood vessels AND holes in blood vessels *allow exit of* phagocytes, therefore, histamine is required for phagocytes to exit). By contrast, a list of terms with few action verbs such as “histamine, blood vessels, phagocytes” would not permit the inference to be made. We argue that inference making during reading is likely to be associated with generation of inferences at post-reading recall. We expect that the more inferences participants verbalize while reading, the more inferences they will make at recall. Inferences at recall that repeat inferences made while reading are likely to be less effortful (at the time of recall) than inferences that have to be drawn at recall. We therefore reasoned that more inferences during reading will be associated with more inferences at recall. This hypothesis is supported by prior research (e.g., Author et al., 2010).

Method

In order to test these hypotheses, we collected think-aloud protocols (Fox, Ericsson, & Best, 2011) from undergraduate students from one large university who had completed a biology course designed for science majors. Participants read seven brief (257 to 387 words), illustrated passages and verbalized their thinking while engaged in reading. Once participants completed the think-aloud, they completed a short demographic questionnaire, followed by a surprise memory task. The memory task required participants to write down everything they remembered (without the texts visible). The transcribed think-alouds (during reading) and recall essays (after reading) were then coded for action verbs, explicit causal terms, and for inferences. These coded data were used to test the hypotheses posed above.

Participants and compensation. Participants were 86 undergraduate students from a large Midwestern university in the United States of America who had completed an introductory organismal and evolutionary biology course designed for biology, biochemistry, bioengineering, and other biology-related STEM majors. Participants had a mean age of 20.0 years ($SD = 1.2$), and had taken the biology course an average of 2.6 ($SD = 1.3$) semesters before the data were collected in Fall 2016. There were 64 (74%) female participants, and 44 (51%) identified as White, 32 (37%) as Asian, 6 (7%) as Hispanic, and 4 (5%) as other or multiple races. Eighteen (21%) were first-generation college students, defined as neither parent with a Bachelor's degree. They were very high-achieving, with mean ACT reading comprehension scores at the 92nd percentile and mean ACT mathematics scores at the 97th percentile. Participants were paid $35 for their time.

Stimuli/Illustrated passages. We presented participants with seven illustrated passages from a biology textbook they would not be familiar with (Sadava, Hillis, Heller, & Berenbaum, 2012). All seven passages focused on different aspects of the immune system. Each passage was presented individually on a computer in the same order as in the original textbook, but with some intervening text edited out. The passages were selected to be relatively brief (257 to 387 words, *M* = 307.2, *SD* = 46.7, including words in the illustrations) and to be comprehensible if read in sequence. The number of instances of action verbs per passage varied from 11 (structure of antibody molecules) to 45 (inflammation as a response to infection; see Table 1).

Insert Table 1

about here

Procedure. Data were collected in an individual laboratory session, beginning with obtaining signed informed consent for each participant as well as agreement to be audiotaped. We started the audio recording, had participants log into a study-specific Blackboard research site, and explained the think-aloud procedure. The directions were modified from Ericsson and Simon (1993) and Authors (2004, 2005, 2010) and comprised the following text "I want you to say out loud EVERYTHING that you are thinking as you are reading the text, this includes reading the text out loud. I would like you to talk out loud CONSTANTLY from the time you begin reading each passage. Say everything that goes through your mind, even if you think it seems irrelevant. I don't want you to try to plan out what you say or try to explain your thoughts, but verbalize your thoughts as they occur. Just act as if you are alone in the room speaking to yourself. It is important that you keep talking. If you are silent for any length of time I will remind you to keep talking aloud. Also, if I can see you doing something, but you are not verbalizing what you are doing, I will ask you to say what you are doing." No specific strategies were ever named or modeled, and there was no practice session. The only two prompts used were "Say what you are thinking" and "Say what you are doing" (e.g., if the participant was taking notes but did not verbalize that he/she was doing that). Participants were also told they had a maximum of 40 minutes to "to learn as much of this material as you can while studying at your usual pace. You have paper, a pen, a pencil, and a highlighter for taking notes, if that is what you usually do when you are studying by yourself from your textbooks. However, I will collect them when you are done reading." No help was provided (e.g., questions were not answered and Internet search was not allowed).

After the 40 minute period, participants were asked to complete a quick demographic survey reporting sex, race, age, and parents' education. Then the computer was closed and participants were asked to move to a different computer in another room for a surprise memory test, where they were asked to type everything they remembered from what they had read (i.e., free recall) into a Word document, with no time limit.

Coding of the stimulus passages. We identified each instance of an action verb used in the passages, since these comprise the action verbs that participants could have picked up from the passage (e.g., bind, make, lyse, dilate, form, split, interact with, etc.; see Table 1). Excluded were state or position verbs such as *is, has, made up of, is found on, lies next to, are connected by*, etc. as well as explicit causal terms (see below). The number of instances of action verbs per passage ranged from 11 to 45 (39, 45, 28, 11, 14, 32, and 19 for passages 1-7 respectively).

We identified each instance of an explicit causal term used in the passages, since these comprise the terms that participants could have picked up from the passage. Explicit causal terms included synonyms for "cause" or "cause and effect" (e.g., because, responsible for, if…then, result in/of, therefore; see Table 2). The number of explicit causal terms per passage ranged from none to 7 (0, 5, 4, 2, 0, 1, and 7 for passages 1-7 respectively).

Insert Table 2

about here

Transcribing and coding think-aloud protocols. We transcribed all audiotapes verbatim, and coded the transcripts for causal terms and action verbs picked up during reading using the same list of words coded in the passage, plus synonyms of those words (e.g., "blows up" for lyses). Each author coded all transcripts (100% per coder x 3 coders), and any disagreements were resolved through discussion. Words were counted as picked up during reading when used in a paraphrase or summary or when taking notes, but not simply during re-reading (see Table 3 for examples of picking up and not picking up causal language). In some cases, the same action verb was used repeatedly in the original passage (CD 14 *activates* TLR, which *activates* the protein kinase cascade), in which case each instance from the during-reading transcript was counted separately. However, if a participant had already picked up a specific instance of an action verb during reading—for example, paraphrasing after reading a few sentences—and later used the same verb when making a summary of those paraphrases after reaching the end of that passage, it was not counted again.

Insert Table 3

about here

We also coded for 5 types of inferences during reading and at free recall: local and global inferences, knowledge elaboration using information from earlier in the text or from before the study began (i.e., not in the text), and hypotheses (Author et al., 2010; see Table 4 for examples). Local inferences were coded for logical conclusions made between two adjacent sentences, and were coded as accurate or inaccurate (i.e., an accurate conclusion or an inaccurate conclusion). Global inferences were coded for logical conclusions made between two non-adjacent sentences within the same passage (also accurate or inaccurate). Knowledge elaborations were coded for logical conclusions between information in the text and information from prior knowledge (likewise, accurate or inaccurate). Two types were coded, one for elaborations that draw on knowledge from an earlier text passage, and one for elaborations that draw on knowledge never mentioned in the passages (i.e., the participant's prior knowledge held before beginning the study). The last type of inference we coded for was hypotheses—predictions about how a process works, based on reasoning from information in the text. Each think-aloud protocol was coded separately by the first and last authors.

Insert Table 4

about here

Participants completed from 4 to 7 passages in their allotted time of 40 minutes. To control for different amounts of text read and the number of causal verbs in each text, we divided the number of action verbs verbalized while completing the think-aloud task by the total number of action verbs contained in the texts read by the participant.[1] For example, participant #6 read passages 1 through 5, which collectively had 137 instances of action verbs (see Table 1), and this participant picked up on 26 of those words, for a proportion of 0.19. These proportions were then used in the analyses. For explicit causal terms and for inferences, we divided the total number verbalized by the number of passages read. These proportions were then used in the analyses.

Scoring free recalls. The typed post-reading free recall passages were coded for instances of explicit causal terms, action verbs, and for inferences not present in the original text, using the same coding described above. Explicit causal terms and action verbs were coded as noted above. Inferences were coded for novel conclusions not present in the text (See Table 4 for examples). These inferences included bridging inferences, knowledge elaborations, and hypotheses as defined above. Each recall was coded separately by the second and last authors (100% coding x 2 coders), and disagreements were resolved by discussion. All three counts—explicit causal terms, action verbs, and inferences—were divided by the number of passages read.

Data Analysis. We conducted the three regressions described above using Ordinary Least Squares linear regression in SPSS Ver. 25 (IBM, Inc., 2018). Variables were screened for normality, linearity, and homogeneity of regression and no violations were noted.

Results

Descriptive statistics on the raw counts and transformed proportions (e.g., action verbs picked up as a percentage of all action verbs in the texts read) of all variables during reading and at recall are shown in Table 5. On average during reading participants picked up on 19% of the

[1] Reading more texts was associated with lower recall scores; this suggests that fast readers were racing through the texts but reading quite shallowly. Even though hypothetically one might associate faster readers (more fluent) with better comprehension, the opposite was the case with these participants.

instances of action verbs in each passage, about one explicit causal term per passage, and made one inference about every 4 passages, but with great variability between participants. To give a sense of the difference between lower and higher level think-alouds, in Figure 1 we show coded sections of the think-aloud transcripts from participants #17 and #18 reading from the same paragraph. In these short examples, Participant #17 picked up on 9 of the action verbs from the paragraph, made 1 inference and picked up on no explicit causal terms. By contrast, Participant #18 picked up on 5 of the action verbs from the paragraph, made no inferences and picked up on one explicit causal term.

Insert Figure 1

about here

At recall, participants wrote about 3.5 action verbs and 0.80 explicit causal terms per passage, and made inferences at about the same rate at recall that they had made while reading, again with great variability between participants. The give a sense of the difference between lower- and higher-level recalls, in Figure 2 we show a portion of the coded recall from two participants who read the same passage. Participant #4 included 14 action verbs in the recall, included one inference in the recall, and included one explicit causal terms in the recall. By contrast, Participant #3 included 3 action verbs in the recall and no inferences or explicit causal terms in the recall.

Insert Figure 2

about here

The low mean rate of pick-up of action verbs while reading suggests participants had limited material for forming a mental model, given that much of the material in paraphrases, summaries, and notes was linked only by vague implicit verbs of classification (‘plasma cell, IgE, mast cells’ versus ‘plasma cells *release* IgE which *binds* to mast cells’). Thus yielding a low-quality textbase (Kintsch, 1998), which then interferes with forming a high-quality situation model.

Insert Table 5

about here

All three regressions were significant overall (see Table 6), explaining 19% of the variance in explicit causal language at recall, 60% of the variance in action verbs at recall, and 14% of the variance in inferences at recall.[2] For generation of explicit causal terms (“because”, “cause,” etc.) at recall as the dependent variable, only making inferences during reading ($\beta = .22$, $p = .03$) was a significant predictor (net of the other two predictors in the model). For generation of action verbs at recall as the dependent variable, two predictors were significant, pick up of action verbs during reading ($\beta = .23$, $p = .003$) and having made inferences during reading ($\beta = .65$, $p < .001$), after accounting for explicit causal language picked up during reading. For generation of inferences at recall as the dependent variable, only pick up of action verbs during reading ($\beta = .21$, $p = .04$) was a significant predictor (net of the other two predictors in the model).

[2] Repeating the analyses as a path model in Mplus—which allows us to obtain overall model fit with three correlated dependent variables—showed virtually the same R^2 values, coefficients and significance tests on the coefficients. The model also showed an excellent fit with $\chi2$ (2) = 0.075, $p = .963$, CFI = 1.000 and SRMR = .011.

Insert Table 6

about here

Discussion

A common misunderstanding in science education is that presenting conceptual models to students—for example, in an illustrated text—can lead to the construction of similar mental models in a student's mind (Greca & Moreira, 2000). However, there is no direct relationship between a displayed conceptual model and a student's mental model, as we have shown with the low level of pick up from illustrated text. Conceptual models are external representations developed by experts to facilitate comprehension of scientific knowledge, whereas mental models are internal representations that learners construct to explain scientific phenomenon or make predictions about how a system works (Nersessian, 1992). Our data provides further support for the notion that simply presenting a conceptual model in the form of an illustrated text to participants does not lead to a similarly structured mental representation for that participant. In fact, we observed that participants picked up only a small portion of the possible action verbs from the text (19%) during reading, indicating that most of their notes, paraphrases, and summaries were focused on structures/substances (e.g., plasma cells, IgE, mast cells, antigen) and included few of the action verbs that link these structures or substances via their function(s) (e.g., plasma cells *produce* IgE, which *binds* to mast cells which *recognize* antigen). They picked up few explicit causal terms during reading (0.71 per passage; e.g., plasma cells produce IgE, which binds to mast cells *in order to* recognize a specific antigen). They also made few inferences during reading, only about 1 per 4 passages.

At recall, participants remembered about four action verbs per passage read, one explicit causal term per passage read, and made slightly more inferences than they had during reading (about one per three passages read). Thus, substantial portions of the recalls consisted of lists of structures or substances and their characteristics—a heavy emphasis on structure—and a small portion consisted of how these structures or substances act on each other and interact with each other—a light emphasis on function. This is consistent with findings from biology education research that students default to memorization when studying and miss structure-function relations, even when instructors emphasize these in their teaching (Kohn, Underwood, & Cooper, 2018).

To construct a well-integrated mental model, learners must identify concepts and relationships in a conceptual model and connect these into a causal model based on general scientific principles (Greca & Moreira, 2000). The strongest results from our analysis concern the mutual influences of action verbs during reading on inferences at recall ($\beta = .22$), and of inferences during reading on action verbs in recall ($\beta = .65$). As hypothesized, failure to pick up on action verbs during reading (i.e., lists of nouns, with vague, implicit verbs of relation) was associated with fewer inferences at recall. This suggests that pickup of action verbs is an important prerequisite for the high-level post-reading inferences that are the desired product of the comprehension process (Kintsch, 1998). When participants made inferences at recall, that was somewhat related to inferences during reading but was ultimately non-significant ($\beta = .17$); inferences at recall were more related to picking up action verbs during reading that are included in those inferences ($\beta = .21$), suggesting that action verb pick up is a prerequisite for inferences.

Interestingly, contrary to prior research (Naumann et al., 2009; Sanchez & Wiley, 2014; Wiley & Voss, 1999; Wiley et al., 2009), the effect of explicit causal language during reading on inferences in recall was not observed in our model. One of the critical differences with the present study is that we separated causal language from action verbs, while previous studies collapsed both types of verbs into one variable. We interpret this finding to mean that it is the action verbs that matter, and this makes sense given that causal language without action verbs is relatively meaningless (e.g., something causes something). In future studies, we suggest that researchers separate causal language from action verbs and analyze the effects of each separately.

Limitations

Our study had a relatively small number of participants, which limits statistical power. Despite this lower than desired power, it is important to note that we still observed statistical significance in the expected relationships. Only one biology topic was used, and the participants came from only one university; results might be different for other topics, other domains, or a different population. Further, the decision to focus on a singular topic, though may limit generalizability across biology topics, does increase the validity of the study. Introductory biology courses tend to cover one topic per week. A third limitation is that we used illustrated text (i.e., the information presented to participants includes both text and diagram), which might introduce unobserved factors such as students' ability to interpret diagrams. In the future study, we can examine the potential effect of text presentation by showing three types of texts to students—text only, diagram only, and illustrated text. In this way, we can understand the possible influence of modality on students' pick-up of causal language.

Conclusion

When participants simply verbally list "steps" from an illustrated passage while reading (e.g., antigen, CD 14, toll-like receptor, protein kinase cascade), it appears they are missing critical causal connections in the action verbs: the antigen *binds* to the CD14, which *activates* the toll-like receptor which *starts* the protein kinase cascade. Thus, it is not enough for action verbs, explicit causal language and causal arrows to be present in illustrated text, readers must pick up on the importance of these actions. We know this because those who skipped over the action verbs had more poorly integrated mental models as indexed by inferences in their free recalls, even after accounting for inferences made during the reading process. Pick up of action verbs during recall was also associated with more explicit causal language at recall. Furthermore, even when participants made inferences while learning, sometimes they did not remember them at recall but they did use more action verbs at recall.

Together, these findings suggest that prompting students to include action verbs when paraphrasing or especially when summarizing or taking notes on illustrated passages might prompt more sophisticated mental model formation. Such prompting could start in high school or even middle school science instruction, allowing for students to fully develop these skills long before they begin undergraduate biology. Action verbs are easy to identify in text, especially compared to explicit causal language (Barriere, 2001), making action verbs an easy target for instruction. This also suggests that, when possible, instructors could carefully choose texts that highlight structures, functions, and the structure-function relation, rather than ones that mostly focus on structures.

References

Author et al., 2010. Peer-reviewed, published journal article.

Barrière, C. (2001). Investigating the causal relation in informative texts. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, *7*(2), 135-154.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*(1), 182-197.

Cho, Y. H., & Jonassen, D. H. (2012). Learning by self-explaining causal diagrams in high-school biology. *Asia Pacific Education Review*, *13*(1), 171-184.

College Board. (2015). *AP Biology course and exam description.* NY, NY: College Board.

College Board. (2018). *AP Biology exam practice.* https://apstudent.collegeboard.org/apcourse/ap-biology/exam-practice

Ericsson, A., & Simon, H. (1993). *Protocol analysis: Verbal reports as data (revised edition).* Cambridge, MA: Bradford books/MIT Press.

Graesser, A., & Britton, B. K. (1996). Five metaphors for text understanding. In B. K. Britton & A. Graesser (Eds.), *Models of understanding text* (pp. 341-351). Mahwah, NJ: Erlbaum.

Greca, I. M., & Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education, 22*(1), 1-11.

Hu, L. T., & Bentler, P. M. (1999).Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.

Kohn, K. P., Underwood, S. M., & Cooper, M. M. (2018). Connecting structure–property and structure–function relationships across the disciplines of chemistry and biology: Exploring student perceptions. *CBE—Life Sciences Education, 17*(2), ar33, 1-15.

Koslowski, B. & Masnick, A. (2010). Causal reasoning and explanation. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development,* 2nd ed. (pp. 377-398). Oxford: Wiley-Blackwell.

Leon, J. A., & Escudero, I. (2015). Understanding causality in science discourse for middle and high school students. Summary task as a strategy for improving comprehension. In K. L. Santi & D. K. Reed (Eds), *Improving reading comprehension of middle and high school students* (pp. 75-98). NY, NY: Springer.

McMaster, K. L., van den Broek, P., Espin, C. A., Pinto, V., Janda, B., Lam, E., ... & van Boekel, M. (2015). Developing a reading comprehension intervention: Translating cognitive theory to educational practice. *Contemporary Educational Psychology*, *40*, 28-40.

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

Montanero, M., & Lucero, M. (2011). Causal discourse and the teaching of history: How do teachers explain historical causality? *Instructional Science, 39*, 109–136. doi: 10.1007/s11251-009-9112-y

Naumann, A. B., Wechsung, I., & Krems, J. F. (2009). How to support learning from multiple hypertext sources. *Behavior Research Methods*, *41*(3), 639-646.

Nersessian, N. J. (1992).How do scientists think? Capturing the dynamics of conceptual change in science. *Cognitive Models of Science, 15,* 3-44.

O'Brien, E. J., Cook, A. E., Lorch, R. F., Jr., & Lorch, R. F. (Eds.). (2015). *Inferences during Reading.* N. Y., N. Y.: Cambridge University Press.

Royer, J. M., Greene, B. A., & Sinatra, G. M. (1987). The sentence verification technique: A practical procedure for testing comprehension. *Journal of Reading*, *30*(5), 414-422.

Sadava, D. E., Hillis, D. M., Heller, H. C., & Bernbaum, M. (2012). *Life: The science of biology (10th Ed).* Baltimore, MD: W. H. Freeman.

Sanchez, C. A., & Wiley, J. (2014). The role of dynamic spatial ability in geoscience text comprehension. *Learning and Instruction*, *31*, 33-45.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal, 46*(4), 1060-1106.

Table 1

List of action verbs coded in the passages and free recalls

Passage 1
Act as
Activate
Alter
Bind
Break
Change
Come from
Coordinate
Create
Defend
Diverge
Divide
Eat
Grow
Identify
Inactivate
Interact
Kill
Label
Present
Produce
Recognize
Split
Start
Tag
Work

Passage 2
Accelerate
Activate
Adhere
Alter
Attack
Attract
Become
Bind
Come
Consume
Contribute
Dilate
Eat
Engulf
Go
Increase
Inhibit
Interact
Involve
Isolate
Kill
Let (things) into
Make
Move
Produce
Promote
Recruit
Release
Speed
Spread
Stops
Release
See
Widen

Passage 3
Act
Activate
Alter
Bind
Change
Encode
Enter
Fight
Get
Identify
Interact
Make
Produce
Receive
Regulate
Sense
Start
Stimulate
Transcribe

Passage 4
Add
Attach
Bind
Break
Broken down
Clump
Code
Do
Digest
Fit
Force
Form
Help
Hold
Identify
Inactivate
Ingest
Interact
Make
Perform
Tag

Passage 5
Activate
Assist
Attached
Become
Bind
Broken
Bursting
Cloning
Deal with
Displayed
Divide
Encoded
Form
Help
Holding
Interact
Kill
Lyse
Make
Pops
Proliferates
Recognize
Released
Respond
Split up

Passage 6

Assist
Activate
Attach
Attack
Become activated
Bind/bound
Block
Breaks it down
Control (go out of, not be in)
Creating
Destroy
Detect
Eat
Ensure
Explode
Force
Go to (apoptosis)
Help
Keep in check
Kill
Lyse/ing
Maintain
Make
Mutated
Prevent
Produce
Protect
Recognize
Regulate
Remove
Secrete
Self-destruct
Stimulate
Suppress

Passage 7
Act
Bind
Exposed to
Learn
Make
Produce
Released
Respond
Secrete
Signal
Shut down
Take

Note: Synonyms for these verbs were also coded (e.g., “blows up” for lyses).

Table 2

List of explicit causal terms coded in the passages and free recalls

Term
allow
are the reason for
as (causal)
as a result of
because (of)
cause
due to
end up leading to
if…then…
in order to
is due to
is ultimately due to
lead to
result from
result in
so (causal)
so as to
so that
the reason why
thus
two reasons for
unless
when (causal not temporal)

Table 3

Examples of "pickup" of action verbs and causal terms from think-alouds and free recall

Source	Example of pickup	Example of non-pickup
Think-aloud examples		
Paraphrase	*Complement proteins leave the vessels and attract phagocytes.* These complement proteins are the ones that attract [action verb] phagocytes.	*Complement proteins leave the vessels and attract phagocytes.* It also shows complement proteins.
Summarize	*It may bind one of the antigen molecules that is already bound to the first antibody, along with a third antigen molecule.* It is stacking [action verb] antigen on top of antigen.	*It may bind one of the antigen molecules that is already bound to the first antibody, along with a third antigen molecule.* So now looking at the picture makes more sense.
	Binding of these fragment molecules to the receptor sets in motion a cascade of molecular changes that results in a change in the three-dimensional structure of the transcription factor NF-κB…The shape change allows NF-κB to enter the nucleus, bind to the promoters of genes, and activate the transcription of genes encoding defensive proteins. So NF-kB because of [causal] the change [action verb] of shape, …enters [action verb] the nucleus, binds [action verb] to those genes, and then transcribes [action verb] those genes.	*Binding of these fragment molecules to the receptor sets in motion a cascade of molecular changes that results in a change in the three-dimensional structure of the transcription factor NF-κB…The shape change allows NF-κB to enter the nucleus, bind to the promoters of genes, and activate the transcription of genes encoding defensive proteins.* NF-kB is a transcription factor.
Taking notes	*They recognize and bind to nonself substances presented with self MHC molecules on the surfaces of other cells.* Recognize [action verb] and bind [action verb] I'm just taking notes.	*They recognize and bind to nonself substances presented with self MHC molecules on the surfaces of other cells.* And to remind myself what MHC stands for uh, major histocompatibility complex

Source	Example of pickup	Example of non-pickup
Free recall examples		
	Histamine is released [action verb] by mast cells, cause [causal] dilation [action verb] of blood vessels; increases [action verb] temperature of the body; which can kill [action verb] pathogens, as well as speed up [action verb] the immune response.	Mast cells [are] first responders.
	The complexes have a heavy and a light chain…there is also a constant segment and a variable segment, the latter being the source of specificity as [causal] this is where the antigen binds [action verb]	Each the light the heavy had I think a constant and a variation one.
	Mast cells that release [action verb] cytokines that draw [action verb] helper cells to an infected site in order to [causal] clean up the site	-chemical signaling

Note: Reading from the text during think-alouds is shown in *Italics*.

Table 4

Coding scheme for inferences in text and free recall

Code ABBREVIATION	Definition	Example
Hypothesis HYP	Pose a hypothesis about how something might work (the hypothesis is not stated in the text)	"All Ts or maybe B cells have the same region as well. Because those are both glycoproteins."
Inference local INFLOC + for accurate - For inaccurate	Participant makes a conclusion across 2 adjacent sentences	"So basically it prevents any non-self cells from getting in and causing infection" "*Thus the two chains have distinct regions with constant and variable amino acid sequences.* It's kind of implied when it says there's two separate genes."
Inference global INFGLOB + for accurate - For inaccurate	Participant makes a generalization across a large segment of text (not just a summary)	"So the immune response attacks itself. Um so Tregs are important to prevent autoimmunity" "B lymphocytes differentiate to form antibody producing cells"
Knowledge Elaboration Before Study KEBS + for accurate - For inaccurate	Participant adds information not in text + info from text and draws a conclusion	"That seems kind of counter-productive, being leaky while you are trying to hold everything in" "Um, damage, I don't know if that's like triggering apoptosis or something"
Knowledge Elaboration Earlier in Text KEET	Participant adds information read in a previous passage + info from current text and draws a conclusion	"OK, this is kind of reminding me of the CD14 because it's, it also binds to a fragment of something."

Code	Definition	Example
ABBREVIATION		
+ for accurate - For inaccurate		"the receptors on T lymphocytes are smaller than the…on B lymphocytes, but their two polypeptides contain both variable and constant regions. That reminds me so much of immunoglobulins"

Note: Reading from the text during think-alouds is shown in *Italics*.

Table 5. *Descriptive statistics on and intercorrelations among all variables*

	1	2	3	4	5	6	7	8	9	10	11	12
1. # action verbs during reading	—											
2. Action verbs during reading (% of words)	.792**	—										
3. # explicit causal terms during reading	.417**	.440**	—									
4. Explicit causal terms during reading (per passage)	.329**	.501**	.860**	—								
5. # inferences during reading	.283**	.398**	.141	.305**	—							
6. Inferences during reading (per passage)	.182	.401**	.111	.398**	.905**	—						
7. # action verbs at recall	.218*	.340**	.074	.095	.401**	.341**	—					
8. Action verbs at recall (per passage)	.173	.494**	.111	.378**	.595**	.744**	.756**	—				
9. # explicit causal terms at recall	-.080	.107	.067	.055	.045	.021	.637**	.418**	—			
10. Explicit causal terms at recall (per passage)	-.030	.351**	.134	.331**	.231*	.354**	.577**	.727**	.815**	—		
11. # inferences at recall	.090	.181	.181	.122	.088	.057	.340**	.251*	.486**	.426**	—	
12. Inferences at recall (per passage)	.108	.321**	.181	.254*	.225*	.286**	.366**	.492**	.450**	.603**	.913**	—
M	32.79	0.19	3.37	0.71	1.11	0.26	16.79	3.48	3.98	0.80	1.31	0.27
SD	14.86	0.09	2.55	0.62	1.12	0.39	9.48	2.74	3.16	0.75	1.57	0.35
Mdn	32.00	0.18	3.00	0.54	0.84	0.17	16.00	3.00	4.00	0.67	1.00	0.17

Notes: # = number of, Prop. = proportion. * indicates significant at $p < .05$, ** indicates significant at $p < .01$

Table 6

Results of Regressions

DV Predictors	*F*	*p*	*MSE*	*b*	β
Explicit causal language at recall	6.53	.001	0.47		
Explicit causal language during reading				0.18	.15
Action verb pickup during reading				1.54	.19
Inferences made during reading				0.42	.22*
Action verbs at recall	40.85	< .001	3.12		
Explicit causal language during reading				0.01	< .01
Action verb pickup during reading				6.94	.23*
Inferences made during reading				4.59	.65*
Inferences at recall	3.34	.007	0.1.07		
Explicit causal language during reading				0.05	.08
Action verb pickup during reading				0.80	.21*
Inferences made during reading				0.15	.17

Notes: *df* = 3, 82 for all ANOVAs, * indicates significant at $p < .05$

Figure 1

Examples of coded segments from the think-aloud protocols, illustrating higher-level activity during reading vs. lower-level activity during reading

---

Higher-level think-aloud

Interactions of Cells and Chemical Signals Result in Inflammation. Histamine and other signals are released from mast cells to initiate the inflammatory response. The chemical signals associated with inflammation attract the phagocytes that digest the pathogens and damaged cells.
1. Damaged tissues attract mast cells which release histamine, which diffuses into the vessels The damaged tissue is giving off [AV] histamine, and it's going to diffuse [AV] into the vessel below it
2. Histamine causes the vessels to dilate and become leaky; so the splinter is out and it looks like there's a lot more histamine in the tissue. The mast cell is gonna release [AV] histamine. So my blood vessel has expanded [AV] and it's becoming leaky [AV] *complement proteins leave the vessels and attract phagocytes.* They're attracting [AV] phagocytes
3. Blood plasma and phagocytes move into infected tissue from the vessels. a phagocyte coming in [AV]
4. Phagocytes engulf bacteria and dead cells. scooping up [AV] all the bad things
5. Histamine and complement signaling cease; phagocytes are no longer attracted. Looks like they're moving [AV] from the vessel. Well I guess the histamine activates something in the blood and it attracts these phagocytes which come and widen the vessel, shimmy through the tissue [INF]
6. A growth factor from platelets stimulates endothelial cell division, healing the wound. The wound is completely healed.

Lower-level think-aloud

Interactions of Cells and Chemical Signals Result in Inflammation. So it's just explaining everything I just read previously so I'm just skimming over the words in the description of the picture *Histamine and other signals are released from mast cells to initiate the inflammatory response. The chemical signals associated with inflammation attract the phagocytes that digest the pathogens and damaged cells* but then there's three different pictures and they're labeled one through six. It kinda shows the splinter and then bacteria produced by [AV] the splinter goes into [AV] the skin. You see that the *1. Damaged tissues attract mast cells which release histamine, which diffuses into the vessels* So those get attracted [AV] and they also diffuse into [AV] the vessel which [ECL] signals [AV] everything else such as like the histamine.
2. Histamine causes the vessels to dilate and become leaky; complement proteins leave the vessels and attract phagocytes.
3. Blood plasma and phagocytes move into infected tissue from the vessels.
4. Phagocytes engulf bacteria and dead cells.
5. Histamine and complement signaling cease; phagocytes are no longer attracted.
6. A growth factor from platelets stimulates endothelial cell division, healing the wound.
I'm done.

Note: Reading from text is shown in Italics; regular text indicates learner verbalizations. Codes attached to think-aloud segments are shown in brackets, such as Action Verb [AV], Inference [INF], and Explicit Causal Language [ECL]

Figure 2

Examples of coded segments from the recalls, illustrating higher-level recall vs. lower-level recall

Higher-level recall

1. The fungal or bacterial cell fragment interacts [AV] with or binds to CD14, then CD14 binds [AV] to the toll-like receptor
2. This catalyzes [AV] a protein kinase cascade
3. This protein kinase cascade leads to a change [AV] in the 3D structure of the transcription factor NF-kB

a. Before NF-kB was in a conformation that prevented it from being able to enter [AV] the nucleus but this conformational change [AV] allows [ECL] for it to enter [AV]

4. The NF-kB enters [AV] the nucleus
5. NF-kB binds [AV] the promoter of the gene that it will help to transcribe [AV] (one of those 40 genes mentioned earlier)
6. With the help of NF-kB, RNA polymerase can be recruited [AV] and the gene will be transcribed [AV]
7. The protein product will then go out [AV] into the cell and combat [AV] the pathogen [INF]

Lower-level recall

Antibodies come from immunoglobulin family
Antibodies have 1 variable region and 1 constant region
Can have primary antibodies, secondary antibodies, etc and they can all form [AV] in a cluster (e.g. secondary body can bind [AV] partially to primary antibody and partially to a third epitope, etc)
Variable region is what binds [AV] to specific epitope
Is a signal transduction pathway – involves a kinase cascade

Note: codes attached to recall segments are shown in brackets, such as Action Verb [AV], Inference [INF], and Explicit Causal Language [ECL]