Toward a typology of integration: Examining the documents model framework

Alexandra List⁎, Hongcui Du, Ying Wang, Hye Yeon Lee
Dept. of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, United States

ARTICLE INFO

Keywords:
- Multiple texts
- Documents model
- Integration
- Composition task

ABSTRACT

Drawing on the Documents Model Framework, in two studies we investigate the types of multiple text models (e.g., mush model, separate representations model, documents model) reflected in students’ written responses, composed based on multiple texts. In among the first studies to holistically code written responses for the type of multiple text models they reflect, we find evidence for all of the multiple text representations specified in the Documents Model. We further examine the types of integration, or cross-textual connections, featured in students’ responses, including evidentiary (i.e., corroborating evidence), thematic (i.e., comparing main ideas across texts), and contextual (i.e., comparing meta-textual information, like author expertise) connections. Finally, we associate the types of multiple text models reflected in students’ written responses with various measures of integration, including the number of discourse connectives and citations included and self-reports of strategy use (Study 1), as well as with integrative utterances reported during a cued think-aloud (Study 2).

1. Introduction

The Common Core Standards for 9th–12th grade English Language Arts ask that students be able to: integrate and evaluate multiple sources of information...in order to address a question or solve a problem (National Assessment Governing Board, 2017). Given its importance, integration has often been examined in prior work as an outcome of multiple text use (Bråten, Stremme, & Britt, 2009; Britt & Aglinskas, 2002; Wiley & Voss, 1999). Nevertheless, the precise nature of integration is not yet well-established. For instance, questions remain regarding how many cross-textual connections students may be expected to form when composing written responses based on multiple texts or the expected specificity of these connections. In two studies, we provide an empirical examination of how students may integrate information when learning about a topic from multiple textual sources and consider the conditions under which integration may be more or less likely to occur.

1.1. Integration

Perfetti, Rouet, and Britt (1999) suggest that connection building, or integration, is core to what distinguishes multiple text processing from the comprehension of single texts, which does not require cross-textual connections to be formed. Models of multiple text comprehension and integration have also been proposed, most prominently the Documents Model of Multiple Texts (DM, Britt, Perfetti, Sandak, & Rouet, 1999; Perfetti et al., 1999). In the Documents Model, Britt et al. (1999) argue that students make sense of multiple texts by constructing two mental representations: the integrated mental model and the inter-text model. The integrated mental model may be considered to be a content-focused cross-textual model, capturing the common topic or issue discussed across texts. The inter-text model is a structural representation of the relations among texts, mapping whether various sources agree, disagree, or complement one another (i.e., source-source links). The inter-text model further maps the relations between specific information and the sources it comes from, through the formation of source-content links.

In forming an overarching representation of multiple texts, then, Britt et al. (1999) propose that students connect their integrated mental model with their inter-text model, to varying degrees of effectiveness. Specifically, these two sub-models may be unified in at least four ways, reflecting differences in the successful integration of multiple texts. First, in a mush model, students are able to form an integrated mental model, or a unified understanding of the content or overarching situation described across texts, but neglect inter-text model development, failing to track which sources information comes from. Because students fail to tag information according to its source of origin, in the mush model, learners are unable to successfully determine the relative...
trustworthiness of different pieces of information or to satisfactorily reconcile conflicts or disagreements between texts, when they arise. As a contrast, in the separate representations model, information is associated with its source of origin in the inter-text model, but integration is poor, meaning that students fail to form a unified picture of the issue introduced across texts in their integrated mental model. While students may adequately understand individual texts, separate from one another, they fail to integrate the information presented across texts in a meaningful way.

Third, the tag-all model is seemingly an ideal representation of multiple texts, wherein students both successfully assimilate each piece of information into their integrated mental model and associate it with its source of origin. Nevertheless, the tag-all model has been deemed to be an expert-level model, too demanding for novice learners to construct effectively. Finally, the documents model of multiple texts is held as the preferable model of multiple text representation for novice learners to develop. The documents model is a more streamlined version of the tag-all model, with important information, appearing across texts, emphasized and irrelevant information discarded. Like the tag-all model, the documents model is both successful in aggregating information across texts, in the integrated mental model, and in tagging information according to source of origin, in the inter-text model.

Looking across the four possible representations that students may construct, only the tag-all, expert-level model, and the documents model, for novices, may be considered to be true reflections of multiple text integration. In this paper, across two studies conducted with a novice sample, we examine students’ documents model formation. Moreover, we examine the extent to which students form documents models when presented with multiple texts and consider the degree and types of integration evidenced in students’ writing.

### 1.2. Assessing integration

Students’ integration of multiple texts has been assessed in two primary ways. The first has asked students to complete researcher-constructed measures of multiple text integration; the second has identified instances of integration in students’ writing. Following a more overt approach to integration assessment, Bråten et al. (2009) asked participants to complete an intertextual inference verification task (Inter-IVT), an integration measure developed based on prior work examining single text comprehension (Royer, Carlo, Dufresne, & Mestre, 1996; Wiley & Voss, 1999). This task asked students to endorse items, created by combining information from two or more texts into a single statement, as either reasonable or unreasonable inferences to draw, based on information presented across texts. This type of assessment has been used in several studies (Braasch, Bråten, Strømso, & Anmarkrud, 2014; Bråten & Strømso, 2010; Bråten, Strømso, & Samuelstuen, 2008; Gil, Bråten, Vidal-Abaca, & Strømso, 2010a, 2010b; Strømso, Bråten, & Britt, 2010) and has determined students to be only moderately effective in source integration. Consistent with other studies, Bråten et al. (2009) found students’ average scores on the Inter-IVT to be 13.48, out of a possible 20 items, indicating that students endorsed only 67% of integrated items correctly.

Using a less directed approach to assessment, Wiley and Voss (1999) examined students’ writing for evidence of multiple text integration. Specifically, they assessed integration in two ways. First, they classified each sentence in students’ responses as added (i.e., containing only novel, non-text-based information), borrowed (i.e., taken directly or paraphrased from textual material), or transformed (i.e., combining information in text with new information or with information from another source), with transformed sentences considered to correspond to integration. Second, they tallied the number of connective terms (e.g., at the same time as, then, with) and causal connectors (due to, because of) used in students’ writing, with these terms used to signify relationships across texts, indicating integration. Elsewhere, the number of citations included in students’ responses has been used as a measure of integration, as has the number of “switches” between texts (i.e., discussing Text 1, then Text 2, then Text 1; Britt & Sommer, 2004). Citation count has been considered to correspond to the formation of source-content links, as a part of documents model construction. Switches between sources have been thought to demonstrate integration by reflecting students’ flexible connecting of content across texts.

Across these approaches to assessing integration, the emphasis has been on capturing how students form connections across texts as well as the number of connections formed (i.e., number of switches, number of citations; Bråten et al., 2009; Britt & Sommer, 2004; Wiley & Voss, 1999). In two studies, we build on this prior work to also consider the types of connections that students may form when integrating multiple texts. Such connections may include connecting main ideas across texts, contrasting evidence presented, or comparing authors’ credentials – all potential markers of integration. In two studies, we first examine the types of integration reflected in students’ written responses (Study 1) and then associate integration, as captured through students’ writing, with integration, as reflected in students’ utterances produced during a cued think-aloud (Study 2).

### 2. Study 1

We had four goals in Study 1. First, we used Britt et al. (1999) Documents Model Framework to characterize the types of multiple text models reflected in students’ written responses, as well as to taxonomize the types of integration demonstrated (RQ1). We further explored our coding of the types of multiple text models identified in students’ responses, by examining the association between these models and two metrics of integration (i.e., discourse connectives, citation use), commonly used in prior research (Britt & Aglinskas, 2002; Britt & Sommer; 2004; Wiley & Voss, 1999; RQ2). We expected the number of discourse connectives included in students’ written responses to correspond to the degree of content integration evidenced in the mush models and documents models constructed (i.e., reflecting integrated mental model construction). We also expected the number of citations in students’ responses to reflect the formation of source-content links, emblematic of separate representations and documents model construction (i.e., reflecting inter-text model development).

#### 2.1. Role of task in integration

Moreover, in Study 1, we examined students’ integration across two different task conditions (i.e., when students were asked either to compose an argument or to write a research report for policy makers, RQ3). Differences in task type have been found to be differentially facilitative of integration (Bråten & Strømso, 2009; Gil, Bråten, Vidal-Abaca, & Strømso, 2010a, 2010b; Kobayashi, 2009). For instance, Wiley and Voss (1999) compared integration when students were asked to use multiple texts to compose an argument or to write a narrative, summary, or explanatory essay. They found argument tasks to improve integration relative to other task assignments. Wiley and Voss (1999) attributed these differences to argument tasks requiring that students personally connect with and transform content, while other tasks (e.g., writing a summary) only required the reproduction of information. They described these differences in task as reflecting a distinction between knowledge-telling and knowledge-transforming. Elsewhere, List and Alexander (2019) have conceptualized different multiple text tasks to vary according to the degree of personalization, or provision of personal opinion or interpretation, they require and their comprehensiveness, or expected coverage of information in texts.

In this study, we examined integration in students’ written responses under two different task conditions. Specifically, we asked students either to develop an argument or to compose a research report for policy makers on the topic of overpopulation. Drawing on work by Wiley and Voss (1999) we included an argument task, expecting it to foster the transformation and integration of information across texts, and a research task, intended to prompt the relay of information to an
external audience, with more limited transformation and integration evidenced, although not to the same limited extent as may be expected of a summary task. As suggested by List and Alexander (2019), we conceptualized the argument task as comparatively high in personalization, but low in comprehensiveness – commonly thought of by students as only requiring the provision of support for their own point of view and the refutation of conflicting perspectives. As a contrast, we expected a research report to be perceived by learners as requiring a more limited degree of personalization but greater comprehensiveness in discussing issues introduced across texts.

### 2.2. Strategies and integration

Finally, we were also interested in the association between integration, as demonstrated in students’ written responses, and students’ engagement in cross-textual strategy use, as assessed via a self-report questionnaire. Specifically, we administered Bråten and Strømsø (2011) Multiple Text Strategy Inventory (MTSI). The MTSI identifies two groups of strategies associated with multiple text use. These are strategies associated with information accumulation (i.e., gathering as much information as possible from available sources) and cross-textual elaboration (i.e., drawing connections or identifying similarities and differences across texts). We expected reports of cross-textual elaborative strategy use to be associated with a greater degree of integration in students’ written responses (Bråten & Strømsø, 2011). Moreover, we expected the argument task, relative to the research report, to foster integration as manifest not only in students’ written responses, but also in their reports of cross-textual elaborative strategy use.

We had the following research questions:

1. To what extent do students’ written responses conform to the models of multiple texts described in the Documents Model Framework? What types of integration are reflected in students’ written responses?
2. What is the degree of association between the type of multiple text models featured in students’ written responses and the number of discourse connectives and the number of citations included?
3. To what extent does integration differ across task types (i.e., arguments vis-a-vis research reports)?
4. What is the association between the degree of integration in students’ written responses and students’ reports of strategy use, as assessed by the Multiple Text Strategy Inventory?

### 3. Methods

#### 3.1. Participants

Participants were 143 undergraduate students (age: M = 20.20, SD = 1.67) at a mid-size university in the Mid-Western United States. The sample was majority female (female: 71.33%, n = 102; male: 27.27%, n = 39) and majority White (67.83%, n = 97). Other students represented in the sample included African American/Black (18.88%, n = 27), Hispanic/Latino (4.90%, n = 7), and Asian (2.10%, n = 3), with two students reporting their race/ethnicity as biracial or multiracial (1.40%) and three reporting their race/ethnicity as Other (2.10%). The sample represented a variety of majors, primarily in the social and natural sciences, and class standings. Specifically, the sample was 24.48% freshmen (n = 35), 33.57% sophomores (n = 48), 25.17% juniors (n = 36), and 15.38% seniors (n = 22). Two students declined to provide demographic information.

### 3.2. Procedures

The study consisted of three parts. First, we asked students to complete a host of individual difference measures, including assessments of prior knowledge, interest, and attitudes. Then, we asked participants to complete a multiple text task, involving researching a complex topic (i.e., threats of and solutions to overpopulation) using a library of six digital texts and composing a written response based on information in texts. For the multiple text task, we randomly assigned participants to one of two task conditions, asking them to either compose an argument or a research report on the threats of and solution to overpopulation.

Following written response composition, we asked participants to complete a variety of post-task assessments, including reporting their strategy use via the Multiple Text Strategy Inventory (Bråten & Strømsø, 2011). Students completed all study components during a single study session, taking place in a computer lab and lasting approximately 1 hour.

#### 3.3. Measures

##### 3.3.1. Individual difference measures

We collected a variety of individual difference measures, including assessments of prior knowledge, interest, and attitudes. We controlled for prior knowledge in all analyses. We include analyses, with all individual difference measures collected in Appendix A. However, we did not find individual difference measures to be associated with performance. In Table 1 we include mean values on individual difference measures in association with the types of multiple text models that students constructed.

##### 3.3.1.1. Prior knowledge

We assessed prior knowledge via a term identification measure. Specifically, we asked participants to define seven terms related to the topic of the task (i.e., overpopulation) and included in the texts provided. Moreover, we instructed participants that: The study will ask you to research and write an argument/research report about overpopulation. To start off, please define each term related to overpopulation. If you don’t know the definition of a term, please write N/A.

Sample terms included: population bomb, high-yield crops, fertility rate, and earth’s carrying capacity. We scored students’ responses to the prior knowledge measure as correct or incorrect, with total prior knowledge scores ranging from zero to seven. Participants had an average prior knowledge score of 3.07 (SD = 1.50), indicating that students constituted a low knowledge sample. Two raters scored 29 students’ responses (20.28%), with Cohen’s kappa equal to 0.76, indicating moderate to strong agreement (exact agreement: 88.18%).

##### 3.3.2. Multiple text task

The multiple text task consisted of two parts, the research phase and the response composition phase.

### Table 1

<table>
<thead>
<tr>
<th>Multiple Text Model</th>
<th>Prior Knowledge M (SD)</th>
<th>Interest M (SD)</th>
<th>Attitudes M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mush Model</td>
<td>3.26 (SD = 1.53)</td>
<td>4.03 (SD = 1.17)</td>
<td>4.71 (SD = 1.00)</td>
</tr>
<tr>
<td>Separate Representation Model without Citations</td>
<td>2.84 (SD = 1.54)</td>
<td>3.51 (SD = 1.20)</td>
<td>4.87 (SD = 0.95)</td>
</tr>
<tr>
<td>Separate Representation Model with Citations</td>
<td>2.76 (SD = 1.30)</td>
<td>3.96 (SD = 0.88)</td>
<td>4.45 (SD = 0.84)</td>
</tr>
<tr>
<td>Documents Model</td>
<td>3.35 (SD = 1.51)</td>
<td>4.21 (SD = 0.92)</td>
<td>4.94 (SD = 0.88)</td>
</tr>
<tr>
<td>Total</td>
<td>3.06 (SD = 1.51)</td>
<td>4.03 (SD = 1.03)</td>
<td>4.73 (SD = 0.91)</td>
</tr>
</tbody>
</table>

Note: No significant differences across multiple text models were found in students’ prior knowledge (p = 0.21), interest (p = 0.08), or attitudes (p = 0.07).
3.3.2.1. Research phase. During the research phase, we presented students with a library of six digital texts and asked them to use these either to compose an argument or a research report. We randomly assigned students to one of two task conditions (i.e., argument and research report). Specifically, we asked students to: Write an argument/ write a research report for policy makers about the threats of overpopulation and how these may be most effectively addressed. Across conditions, we then presented participants with a library of six digital texts. We further instructed participants to: Please take notes while you research. You will not be able to refer to the sources while writing your answer. Fig. 1 includes a screenshot of the directions that participants received.

3.3.2.1.1. Texts. We selected six texts to provide students with a variety of complementary and partially conflicting perspectives on the threats of and solutions to overpopulation. For instance, while one text suggested that overpopulation was not a major issue since, in the future, technology would be able to mitigate its threats, another text considered overpopulation to be a serious threat and suggested the development of new technologies for food production, specifically, as a mechanism for dealing with overpopulation. We selected texts from the Room for Debate segment of the New York Times (https://www.nytimes.com/roomfordebate/2015/06/08/is-overpopulation-a-legitimate-threat-to-humanity-and-the-planet and https://www.nytimes.com/roomfordebate/2011/05/04/can-the-planet-support-10-billion-people) and modified them for inclusion in this study. Please see Table 2 for length and readability.

We selected texts that we deemed to be trustworthy and that provided students with information relevant to response composition. We initially presented texts to students by title; however, students could access document information associated with each text. The document information we provided included author name and publisher, as well as author credentials or affiliation. We attributed all texts to expert authors and reputable publishers in the popular press (e.g., BBC World; The Economist). Students could access as many of the six texts as they wanted and were free to revisit texts. A screenshot of the digital library is provided in Fig. 1. We further randomly assigned students to take notes either using paper and pencil or the computer and students were able to take notes during task completion. We did not expect modality of note-taking condition to have an effect on integration and therefore note-taking conditions are collapsed in this study.

3.3.2.2. Response phase. Once students had completed the research phase of the study, we asked them to compose a written response, corresponding to their task condition (i.e., to write an argument or a research report). We scored students’ written responses in three primary ways. First, we coded students’ responses holistically for the type of multiple text model they represented, as per the Documents Model Framework (i.e., mush model, separate representations model, documents model, tag-all model, Britt et al., 1999; Perfetti et al., 1999). Second, we identified the specific statements in students’ responses indicating integration and coded these for the type of integration they reflected. Third, consistent with prior work, we totaled the number of connectives and citations included in students’ written responses, as proxy measures of integration (Britt & Aglinskas, 2002; Wiley & Voss, 1999).

3.3.2.2.1. Documents model. We first coded students’ responses for the type of multiple text representation (i.e., multiple text model) they reflected. We coded responses as representing a mush model if they discussed information, presented across texts, in an integrated fashion, but exhibited no sourcing or a lack of attribution of information to its document of origin (Britt et al., 1999; Perfetti et al., 1999). Second, we identified the specific statements in students’ responses indicating integration and coded these for the type of integration they reflected. Third, consistent with prior work, we totaled the number of connectives and citations included in students’ written responses, as proxy measures of integration (Britt & Aglinskas, 2002; Wiley & Voss, 1999).

Table 2
Readability across texts.

<table>
<thead>
<tr>
<th>Source</th>
<th>Readability</th>
<th>Population is a Major Challenge</th>
<th>Word Count</th>
<th># of Sentences</th>
<th>Flesch Reading Ease</th>
<th>Flesch-Kincaid Grade Level</th>
<th>Coh-Metrix L2 Readability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overconsumption is a Grave Threat to Humanity</td>
<td>No</td>
<td></td>
<td>272</td>
<td>21</td>
<td>44.8</td>
<td>10.5</td>
<td>11.87</td>
</tr>
<tr>
<td>Violent Side Effects of High Fertility Rates Technology and Population</td>
<td>Yes</td>
<td>No</td>
<td>270</td>
<td>15</td>
<td>38.3</td>
<td>12.1</td>
<td>11.90</td>
</tr>
<tr>
<td>Build a Less Wasteful Economy</td>
<td>Yes</td>
<td></td>
<td>222</td>
<td>10</td>
<td>23.8</td>
<td>14.8</td>
<td>11.21</td>
</tr>
<tr>
<td>Empowered Women for the Health of the Planet</td>
<td>Yes</td>
<td></td>
<td>274</td>
<td>13</td>
<td>46.7</td>
<td>13.2</td>
<td>15.35</td>
</tr>
<tr>
<td>More Efficient Food Production</td>
<td>No</td>
<td></td>
<td>251</td>
<td>8</td>
<td>25.6</td>
<td>17.4</td>
<td>14.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>269</td>
<td>14</td>
<td>35.1</td>
<td>13.1</td>
<td>11.27</td>
</tr>
</tbody>
</table>
Table 3
Integration taxonomy.

<table>
<thead>
<tr>
<th>Integration Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explicit Integration</strong></td>
<td>Even though we do not hold the largest population in the world, compared to Asia's population estimated to reach five billion by 2050 or the 25% of all children in the world found in Africa, the US still manages to use the most (Build a Less Wasteful Economy, The Violent Side Effects of High Fertility). [Complementary]</td>
</tr>
<tr>
<td><strong>Implicit Integration</strong></td>
<td>If we were to cut down tremendously on the amount of meat consumed in the west and moved toward a primarily vegetarian diet, these issues would be drastically diminished. If this were to happen, the land formerly used for grazing animals and growing food to feed them could be used for more effective things like growing food for people to eat. (Overconsumption is a Real Threat to Humanity). 33% of the Earth's surface is used for food and is still expanding (More Efficient Food Production). [Complementary]</td>
</tr>
<tr>
<td><strong>Thematic Integration</strong></td>
<td>Cascio, writer for the New York Times…indicates that the issue isn't an overpopulation bomb, but rather an &quot;overconsumption bomb.&quot; Jason Clay, author of the article “More efficient food production” also seems to back up this line of thinking. [Complementary] Another provided source stated quite the opposite of Cascio. Jason Clay of Foreign Affairs stated that the United Nations projection shows a troubling trend of increased population and rising consumption on a planet with finite resources. [Conflicting]</td>
</tr>
<tr>
<td><strong>Contextual Integration</strong></td>
<td>People like Barroso and Cascio are focused on the roles women play in population. [Complementary] The best way to address these different viewpoints is to address them separately. (4) I'm not quite sure how I'd approach someone with a political view (Building a Less Wasteful Economy) on overpopulation…Although, if I was addressing someone with an eco-friendly standpoint (Overconsumption) about overpopulation…[Contrasting]</td>
</tr>
</tbody>
</table>

Note: Citations are underlined; connective terms and unifying statements are bolded.

Separate representations responses typically presented information from different sources in a sequential manner, but with no cross-textual integration. A variant of this model was a separate representations model without citations. This model was similar to the separate representations model with citations, in that it presented information sequentially, by source of origin, with no integration evidenced. But this model differed from the separate representations model in that it did not cite sources or connect information back to its document of origin.

We placed responses into the documents model category if they exhibited some degree of integration, forming connections between two or more texts, and explicitly referred to sources through citation. Finally, much like the documents model, we coded responses as corresponding to a tag-all model of multiple texts, when they explicitly cited and integrated texts. However, we distinguished these responses from a documents model by the volume of integration they evidenced; these responses comprehensively connected all six texts from the digital library to one another. Given the expert-level nature of tag-all models, we were not surprised that only one student response was classified as representing a tag-all model in our sample.

Two raters coded 30 students’ responses (20.98%) for the type of multiple text model manifest, with Cohen’s kappa equal to 0.87, indicating strong inter-rater agreement. We resolved all disagreements through discussion.

3.3.2.2.2. Integrative statements. Once student responses were identified as featuring multiple text integration (i.e., mash model, documents model, tag-all model) or not (i.e., separate representations model, with and without citations), we coded for the specific nature of the integration featured in students’ responses. We first identified responses as including instances of explicit or implicit integration. Explicit integration was evidenced in responses that included a single statement expressly linking two or more texts to one another. We further identified such explicitly integrative statements by their inclusion of a conjunctive or connecting term to signify a relation between two or more texts (e.g., agrees, also, disagrees, conflicts).

For instance, we considered a statement like: People like Barroso and Cascio are focused on the roles women play in population, to demonstrate explicit integration as it drew together arguments presented by two different authors, using the connective term and. Implicit integration was demonstrated when students produced a single, unifying statement to summarize content across texts and then substantiated that statement with specific references to sources that the statement characterized. In this way, while texts were not explicitly linked to one another, they were implicitly connected through their relevance to and positioning in relation to a particular unifying statement. An example of implicit integration included a student writing:

“The biggest threat of overpopulation is overconsumption. Jamias Cascio wrote in “Overconsumption is a Grave Threat” that overconsumption is when depletion of necessitiessuch as soil, water, and other life-supporting resources happen. Overconsumption is so bad that Jason Clay says that we will need three Earths to support human activities in 2100 if our activities stay at the same level of growth we have now (More Efficient Food Production).”

In this response, the student is implicitly integrating arguments made by two separate authors based on their joint reference to the threat of overconsumption. However, these two texts are connected implicitly, only through their positioning, rather than explicitly, via a connective term. Two raters again scored 30 students’ responses for the type of integration included. Cohen’s kappa inter-rater agreement was 0.97 for the number of instances of explicit integration identified and 0.83 for the number of instances of implicit integration, both indicating strong agreement.

After we coded responses as demonstrating explicit or implicit integration, we further coded integrative statements for their degree of specificity or for the level of connection between texts that they demonstrated. We found integration, be it explicit or implicit in nature, to reflect one of three levels of cross-textual linking. First, we found evidentiary integration to be evidenced when specific information or data were corroborated and linked across texts. Second, we considered thematic integration to be reflected when main ideas were connected across texts. Finally, we found that contextual integration occurred when aspects of the texts, beyond the content they included, were compared or otherwise linked. Contextual integration included the comparisons of author credentials or the linking of disciplinary perspectives across texts.

Considering whether integration was explicit or implicit and whether it was evidentiary, thematic, or contextual in nature resulted in our developing a taxonomy of the ways in which multiple texts may be integrated. This taxonomy is presented in Table 3. Two raters scored each response for instances of evidentiary, thematic, or contextual
integration identified. Cohen's kappa inter-rater agreement was 0.87 for instances of evidentiary integration, 0.93 for instances of thematic integration, and 1.00 for instances of contextual integration, for 30 student responses.

We first examined responses for their inclusion either of integrative statements (i.e., connecting two or more texts in an explicit or implicit fashion) or citations. Once the integrative statements and citations in a response were identified, we categorized written responses according to the type of multiple text model they reflected. Finally, we identified the specific integrative statements in students’ responses as evidentiary, thematic, or contextual in nature.

3.3.2.2.3. Discourse connectives. Consistent with prior work (Wiley & Voss, 1999), we also examined the number of discourse connectives included in students’ written responses. Discourse connectives are connection words (e.g., because, although) used to link disparate concepts or ideas in a single statement. Discourse connectives examined in this study were drawn from the Penn Discourse TreeBank (Miltsakaki, Joshi, Prasad, & Webber, 2004; Miltsakaki, Prasad, Joshi, & Webber, 2004).

3.3.2.2.4. Citations. We also examined the number of citations included in students’ written responses. Citations were explicit references to sources, either by title or by author. Students received a point for each non-consecutive reference to texts included in their written responses. Two raters scored 30 student responses for the number of citations included, with exact agreement equal to 96.67%. The citations that students included in their written responses represented instances of spontaneous, rather than cued, citation, as we did not instruct students to reference texts in the responses that they composed.

While we considered the number of discourse connectives in students’ written responses to be a measure of integrated mental model formation, or students’ development of content-content links, we used the number of citations in students’ responses to assess inter-text model formation, reflecting source-content link development. As such, we expected the number of discourse connectives, in reflecting integrated mental model formation, and the number of citations, in reflecting inter-text model construction, to jointly correspond to documents models manifest in students’ written responses.

3.3.3. Strategy inventory

Following multiple text task completion, we asked students to complete a variety of post-task assessments, including a strategy inventory. We used Braten and Strømsø (2011) Multiple Text Strategy Inventory (MTSI). Cronbach’s alpha reliability for the five item information accumulation scale was 0.82. Reliability for the ten item cross-textual elaboration scale was 0.80.

4. Results

4.1. Research question 1: Models of multiple texts and integration in written responses

4.1.1. Multiple text models

The first research question examined the types of multiple text
models reflected in students’ written responses as well as the type of integration featured. Mash models, including content integration but no citation, were reflected in 18.88% (n = 27) of student responses. Models reflecting separate representations of content, with no integration, were characteristic of 41.96% of student responses (n = 60). This included 13.29% (n = 19) of responses reflecting the separate representation of content with no citation and 28.67% (n = 41) of responses reflecting a separate representation model of multiple texts, with citation, but with no source-source links evidenced. Finally, 34.27% (n = 49) of responses evidenced integration, indicative of documents model formation. We considered one response to be consistent with a tag-all model. We did not code four responses in the sample because they only included information from a single text. Sample responses are presented in Table 4.

We examined whether the different multiple text models that students constructed in their written responses differed in the amount of integrative statements they included. For these analyses, we excluded the student composing a tag-all model of multiple texts. The ANOVA was overall significant, F(3,132) = 48.35, p < 0.001, η² = 0.52, indicating a large effect (Cohen, 1988). As we expected, students constructing documents models (M = 2.12, SD = 1.39) and mash models (M = 1.63, SD = 1.01) included significantly more instances of integration in their written responses than did students constructing separate representation models with (M = 0.00, SD = 0.00, p < 0.001) and without (M = 0.00, SD = 0.00, p < 0.001) citations. However, using Tukey’s post-hoc pair-wise comparisons, we did not find the degree of integration reflected in responses corresponding to a documents model to be significantly greater than in responses reflecting a mash model of multiple texts (p = 0.14). As a note, when we excluded students constructing separate representations models, with or without citations, from analyses, due to their responses definitionally including no integration, no significant differences were identified (p = 0.11).

4.1.2. Types of integration

We were further interested in the types of integration featured in students’ written responses. On the whole, students’ degree of integration was quite limited (M = 1.11, SD = 1.47). Even among students including any integration at all in their written responses, we identified only an average of 2.05 (SD = 1.44) instances of integration, with 51.32% (n = 39) of students including any integration, including only one such instance. In Table 5, we summarize the type of integration included in the various multiple text models reflected in students’ written responses.

4.1.2.1. Explicit and implicit integration. Implicit integration, connecting texts via a general unifying statement (M = 0.82, SD = 1.23), was more common in students’ written responses than was explicit integration (M = 0.29, SD = 0.67), connecting two or more texts via a linking term. Specifically, 18.88% (n = 27) of students included explicit integration in their written responses, while 44.76% (n = 64) of students’ responses featured implicit integration.

4.1.2.2. Evidentiary, thematic, and contextual integration. We further coded students’ explicitly and implicitly integrative statements as reflecting evidentiary, thematic, or contextual integration. The majority of the integration that students demonstrated occurred at the evidentiary (M = 0.55, SD = 0.93) or thematic (M = 0.51, SD = 0.93) level. Contextual integration, reflecting integration based on author information or other document features, occurred least commonly (M = 0.05, SD = 0.30). Indeed, 34.97% (n = 50) included at least once instance of evidentiary integration in their written responses, as compared to 30.07% (n = 43) of students including thematic integration, and 2.80% (n = 4) students including any contextual integration.

4.2. Research question 2: Discourse connectives, citations, and response integration

4.2.1. Discourse connectives

We were interested in whether the type of multiple text model reflected in students’ written responses was associated with differences in the number of discourse connectives that students incorporated into their written responses. Again, we excluded the student constructing a tag-all model of multiple texts. We found the four models to differ significantly in the number of discourse connectives featured in students’ responses, F(3, 132) = 6.37, p < 0.001, η² = 0.13, indicating a medium to large effect. However, post-hoc analyses using Tukey’s HSD only found significant differences between the number of discourse connectives included in students composing mash models (M = 10.22, SD = 3.52) vis-à-vis documents models (M = 14.14, SD = 4.25), p < 0.001.

4.2.2. Citations

We further examined the association between the type of multiple text model reflected in students’ written responses and the number of citations they included. We excluded the student constructing a tag-all model in their written response. We found significant differences, F(3, 132) = 71.23, p < 0.001, η² = 0.62, indicating a large effect. As we expected, significantly more citations were included in students’ responses classified as documents models (M = 4.55, SD = 2.25) and separate representation models, with citations (M = 2.88, SD = 1.25), than were in responses categorized as mash models (M = 0.04, SD = 0.19), and separate representation models, without citations (M = 0.00, SD = 0.00), p < 0.001. More importantly, using Tukey’s post-hoc pair-wise comparisons we found students to include significantly more citations in responses reflecting documents model construction (M = 4.55, SD = 2.25) than in responses reflecting a separate representations model of multiple texts, with citations

Table 5

<table>
<thead>
<tr>
<th>Multiple Text Model</th>
<th>Total Integration</th>
<th>Explicit</th>
<th>Implicit</th>
<th>Evidentiary</th>
<th>Thematic</th>
<th>Contextual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mush Model (n = 27)</td>
<td>M = 1.63 (SD = 1.01)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 1.63 (SD = 1.01)</td>
<td>M = 1.00 (SD = 0.83)</td>
<td>M = 0.63 (SD = 0.84)</td>
<td>M = 0.00 (SD = 0.00)</td>
</tr>
<tr>
<td>Separate Reps. w/o Citations (n = 19)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
</tr>
<tr>
<td>Separate Reps. w/ Citations (n = 41)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
<td>M = 0.00 (SD = 0.00)</td>
</tr>
<tr>
<td>Documents Model (n = 41)</td>
<td>M = 2.12 (SD = 1.39)</td>
<td>M = 0.84 (SD = 0.92)</td>
<td>M = 1.29 (SD = 1.15)</td>
<td>M = 0.94 (SD = 1.03)</td>
<td>M = 1.04 (SD = 1.17)</td>
<td>M = 0.14 (SD = 0.50)</td>
</tr>
<tr>
<td>Tag All Model (n = 49)</td>
<td>M = 8.00</td>
<td>M = 8.00</td>
<td>M = 5.00</td>
<td>M = 3.00</td>
<td>M = 0.00</td>
<td></td>
</tr>
<tr>
<td>Total (n = 137)</td>
<td>M = 1.14 (SD = 1.48)</td>
<td>M = 0.30 (SD = 0.68)</td>
<td>M = 0.84 (SD = 1.24)</td>
<td>M = 0.57 (SD = 0.94)</td>
<td>M = 0.52 (SD = 0.94)</td>
<td>M = 0.05 (SD = 0.30)</td>
</tr>
</tbody>
</table>

Note: Separate Representations models, with and without citations, did not contain any integration.
4.3. Research question 3: Integration across task conditions

We were further interested in the extent to which differences in task assignment (i.e., asking students to write an argument or a research report) was associated with the degree of integration evidenced in students’ written responses. Using a chi-squared test we found a significant association between task condition and the type of multiple text model constructed. \( \chi^2(3) = 7.78, p = 0.05 \), Cramer’s \( V = 0.24 \), indicating a small to medium effect. Based on our examination of standardized residuals, we found that students in the argument condition were disproportionately more likely to construct models reflecting a separate representation of multiple texts, without citation; while students assigned to the research task were disproportionately more likely to engage in mush model construction. We present descriptive information by task assignment in Table 7.

We also examined the association between task condition and the total degree of integration included in students’ written responses. The ANOVA was overall significant, \( F(1, 139) = 4.18, p < 0.05, \eta^2 = 0.03 \), indicating a small to medium effect. Specifically, the students that we assigned to compose the research report included significantly more instances of integration in their written responses (\( M = 1.36, SD = 1.60 \)) than the students who we asked to compose arguments (\( M = 0.86, SD = 1.30 \)).

Finally, we examined the association between task assignments and students’ use of discourse connectives and citations in their written responses. However, ANOVAs comparing discourse connectives (\( p = 0.39 \)) and citation (\( p = 0.70 \)) use across task conditions were not significant.

Table 7

<table>
<thead>
<tr>
<th>Argument (( n = 72 ))</th>
<th>Research Report (( n = 69 ))</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration</td>
<td>( 0.86 (SD = 1.30) )</td>
<td>( 1.36 (SD = 1.60) )</td>
</tr>
<tr>
<td>Discourse Connectives</td>
<td>( 11.79 )</td>
<td>( 12.80 )</td>
</tr>
<tr>
<td></td>
<td>( (SD = 4.02) )</td>
<td>( (SD = 4.48) )</td>
</tr>
<tr>
<td>Citation</td>
<td>( 2.39 (SD = 2.26) )</td>
<td>( 2.55 (SD = 2.64) )</td>
</tr>
</tbody>
</table>

4.4. Research question 4: Integrative strategy use and response integration

Our fourth research question examined the association between integration and students’ scores on the Multiple Text Strategy Inventory. We examined whether there were differences in the extent to which students reported engaging strategies associated with information accumulation and cross-lexical elaboration and the type of multiple text model reflected in their written responses. The degree of information accumulation strategies reported did not differ across the type of multiple text models characterizing students’ written responses, \( p = 0.92 \). However, there were differences in the degree of cross-lexical elaborative strategy use reported across multiple texts models, \( F(3, 133) = 4.05, p < 0.01, \eta^2 = 0.09 \), indicating a medium to large effect. Post-hoc pair-wise comparisons using Tukey’s HSD determined that there was a significant difference in the degree of cross-lexical elaborative strategy use reported between students formulating a separate representations model of multiple texts, with no citations (\( M = 4.76, SD = 0.83 \)), and a documents model (\( M = 5.51, SD = 0.82, p < 0.01 \)) in their written responses.

While the number of integrative statements included in students’ written responses was not significantly associated with students’ reports of information accumulation strategy use (\( p = 0.96 \)), there was a significant association between students’ reports of cross-lexical elaborative strategy use and the total number of integrative statements included in their written responses, \( r(139) = 0.25, p < 0.01 \), indicating a small to medium association.

4.5. Follow-up validation analyses

We used discriminant function analysis to confirm our classification of students’ responses into the multiple text models reflected in the Documents Model Framework. We entered the number of integrative statements, discourse connectives, and citations in students’ written responses as well as students’ scores on the cross-lexical elaboration scale of the MTSI as predictors. We identified two significant discriminant functions, with eigenvalues greater than one. The first explained 59.35% of the variance in multiple text model classification (canonical \( R^2 = 0.63 \)), while the second explained 39.60% of the variance (canonical \( R^2 = 0.53 \)) in students’ response classification. In combination, these discriminant functions significantly differed across students whose responses reflected different models of multiple texts. [Function 1: Wilk’s \( \Lambda = 0.17, \chi^2(12) = 228.26, p < 0.001 \)]. Even when the first function was excluded, the second function was nevertheless still significant [Function 2: Wilk’s \( \Lambda = 0.46, \chi^2(6) = 100.92, p < 0.001 \)]. While citation use loaded significantly onto the first function (\( r = 0.94 \)), the total number of integrative statements loaded significantly onto the second function (\( r = 0.99 \)). Because the cross-loadings for these predictors were fairly small (i.e., \( rs < 0.33 \)) we concluded that both integrative statements and citation use were distinct contributors to response classification. As can be seen in the discriminant function plot in Fig. 2, the first function discriminated multiple text models according to their citation use (i.e., discriminating mush models and separate representation models without citations from separate representations models with citations and documents models). The second function discriminated students’ written responses according to the number of integrative statements included, with mush models and documents models discriminated from separate representations models with and without citations.

5. Discussion

This study contributes to the literature on integration and multiple text use in at least four ways. First, we examine students’ written responses to understand the types of multiple text models that students construct and to identify the extent to which these reflect multiple text integration. Second, we present a taxonomy of the types of integration...
that students may engage in when composing written responses based on multiple texts. Third, we examine a variety of commonly used metrics of response integration (i.e., number of discourse connectives, citations), as well as the MTSI, for their correspondence to the degree of integration reflected in students' written responses. Finally, we examine the role of task in facilitating response integration.

5.1. Research question 1: Models of multiple texts and integration in written responses

We were first interested in examining the extent to which students' written responses manifest the kinds of multiple text models specified in the Documents Model Framework. Although we found students' multiple text integration to be limited, there was evidence for all of the Documents Model typologies introduced by Britt et al. (1999). In particular we found, 34.27% of student responses to reflect documents model construction, or to evidence some degree of integration, albeit limited, alongside the use of citations. Nevertheless, the amount of integration evidenced even among documents models constructed was quite limited (M = 2.14, SD = 1.41).

Focusing more specifically on the types of integration demonstrated in students’ written responses, we identified a number of integration categories. The types of integration featured in students’ written responses differed both in their mode of presentation (i.e., explicit or implicit) and in their level of specificity (i.e., evidentiary, thematic, and contextual). While prior work has not fully considered the types of integrative connections that may be drawn across texts, in this study we provide insights into the range of ways that multiple texts may be linked.

5.2. Research question 2: Discourse connectives, citations, and response integration

Our second research question examined the association between the kinds of multiple text models reflected in students’ written responses and more intermittent metrics of integration, used in prior work. These included the number of discourse connectives in students’ written responses, corresponding to the degree of content integration evidenced, or integrated mental model construction (Wiley et al., 2009); and the number of citations in students’ responses, indicating the formation of source-content links and corresponding to inter-text model formation (Britt & Aglinskas, 2002). As demonstrated in Fig. 3, we expected the four multiple text models coded for in students’ written responses to differ in both the number of discourse connectives featured and in the number of citations included.

The results we found were somewhat mixed with regard to the correspondence between the indicators of interest and the multiple text models that students constructed. In terms of citation use, the results were as predicted. Students composing separate representation models with citations and documents models included significantly more citations in their written responses than students composing either mush models or separate representations models, without citation. Moreover, students constructing documents models included significantly more citations than did students creating separate representations of multiple texts, with citations. This latter finding is notable because constructing a separate representations model may be considered to be a fairly effective means of writing about multiple texts. Such a model is organized, in presenting content from multiple texts, and includes citation use. Nevertheless, separate representations models are limited in their integration of multiple texts, or the formation of source-source links, as reflected in the more limited number of citations used in their composition. Yet, students’ written responses reflecting a documents model...
of multiple texts differed only from mush model responses in the number of discourse connectives included. This may have been the case for two reasons. First, we examined a rather large set of discourse connectives in this study and we did not consider connectives use, as a measure of integration, to be sensitive enough to distinguish between students’ formation of inter-textual links, or connections between texts, verses intra-textual links, within a single text. Moreover, even when constructing documents models, the degree of integration that students exhibited was nevertheless quite limited, overall, potentially resulting in more limited differences in discourse connectives use across the different responses that students constructed. Nevertheless, in future work we hope to carefully look at the role of connectives in linking ideas within a single text or ideas across texts, in the written responses that students compose.

5.3. Research question 3: Integration across task conditions

We further examined the prevalence of multiple text integration, featured in students’ written responses, across task conditions. Based on prior work, we expected the argument task to facilitate integration to a greater extent than the research task (Wiley & Voss, 1999). But, rather, we found that it was the research report that prompted a greater degree of integration in students’ written responses, as compared to the argument task. In particular, the research report was associated with a greater prevalence of mush models constructed, as compared to the argument task, which resulted in a greater number of separate representation models, with no citations, to be developed. It should be noted that mush models, while reflecting integration, did not do so to the same extent as did documents models. Indeed, documents model construction seemed to operate fairly uniformly across task conditions.

We can explain these somewhat unexpected findings in at least three ways. First, prior work examining performance differences across task conditions has found the effectiveness of different task assignments to be moderated by individual difference factors, like prior knowledge (Gil et al., 2010a, 2010b). Second, prior work has suggested that students conceptualize argument tasks differently from research tasks (List, Du, & Wang, submitted for publication). Specifically, it seems to be the case that students consider research reports to require a deeper degree of engagement with multiple texts, than do argument tasks, which students more often consider to require the provision of a personal perspective on texts. As such, when we asked students to compose a research report, rather than an argument, this may have resulted in students exhibiting a greater degree of integration in their written responses, due to research reports stimulating a greater degree of engagement with multiple texts. Connecting these two explanations, it may be the case that while task features improved integration to a limited extent (i.e., such as that exhibited in mush model construction), documents model construction was more contingent on individual differences in integration skills.

Third, it may be that our specifying a particular audience for the research report (i.e., policy makers) resulted in students’ better representation of multiple texts, as compared to the argument task. Indeed, List and Alexander (2019) recently conceptualized argument tasks vis-à-vis research reports as differing, in part, according to the degree of personalization and information selectivity they demand from students. In other words, while an argument task may cue students to include their own points of view or support for their positions to a greater extent, research reports may encourage students to adopt a more balanced perspective on the information presented across texts. Of course, in part, this may be result of students having limited schema for what writing a quality argument entails, namely the reasoned consideration of both sides of a controversial issue. Alternately, it may be that if we provided students with more explicit instructions to integrate, or form connections across texts, this would stimulate documents model construction to the greatest extent.

5.4. Research question 4: Integrative strategy use and response integration

Our fourth research question examined whether multiple text model formation in students’ written responses was associated with multiple text strategy use. Prior work has found students’ reports of information accumulation to be negatively associated with integration, while cross-textual elaborative strategy use was positively associated (Bråten & Stremme, 2011). In this study, no differences in accumulation strategy use were found across the different multiple text models that students constructed; however, we found documents model construction to be associated with a higher degree of cross-textual elaborative strategy use reported. On the one hand, this further validates the cross-textual elaboration dimension of the MTSI as a means of assessing students’ engagement in integration-related strategy use. On the other hand, students rated highly their engagement of strategies associated both with information accumulation and with cross-textual elaboration. Indeed, students indicated that their engagement in cross-textual elaboration was considerably more extensive than we found to be reflected in their written responses. This may have been due to students inflating their degree of strategy engagement, reflecting a limitation of self-report as a means of strategy assessment. Or because the degree of integration that students engaged in during text processing was not able to be translated into the written responses that they composed. To explore this issue further, we carried out Study 2 to examine the relation between students’ integrative strategy use during multiple text task completion and integration, as evidenced in their written responses. In Study 2, in addition to assessing strategy use via the MTSI, we adopted a cued think-aloud protocol to tap integrative strategy use during multiple text task completion.

In Study 1 we focused on integration, as reflected in the written responses that students composed. In Study 2 we were further interested in linking students’ integrative processing, during text use, to the integration demonstrated in the written responses that they composed.

6. Study 2

We carried out Study 2 to validate findings from Study 1. Specifically, we were interested in documenting the association between students’ cross-textual strategy use, as reported during multiple text use, and multiple text model construction, as reflected in the written responses that students composed. In Study 2 we adopted a cued think-aloud protocol to capture the degree of integrative strategy use that students reported during the course of multiple text task completion and to associate such strategy use with the degree of integration evidenced in students’ written responses. We had the following research questions:

1. What types of multiple text models are evidenced in students’ written responses?
2. To what extent do students engage in integrative strategy use during the course of multiple text task completion?
3. What is the association between students’ integrative strategy use, reported during the course of task completion, and multiple text model formation in students’ written responses?
4. What is the association between integrative strategy use, reported during the course of a multiple text task, and students’ strategy use, as reported on the MTSI?

7. Methods

7.1. Participants

Participants were 32 undergraduate students at a mid-size, Mid-Western University (age: $M = 20.90, SD = 4.20$). The sample was majority female (73.33%, $n = 22$; male: 26.67%, $n = 8$) and majority White (75.00%, $n = 24$). The remaining participants reported Black/
African American (12.50%, n = 4) and Asian ethnicity (3.13%, n = 1). The sample represented a variety of majors in the social sciences and a variety of class standings. Specifically, 33.33% (n = 10) were freshmen, 13.33% were sophomores (n = 3), 23.33% were juniors (n = 7) and 30.00% were seniors (n = 9). Two students did not provide demographic information. Due to technical challenges, complete data are only available for 28 participants.

7.2. Procedures

The procedures for Study 2 were identical to those used in Study 1. Specifically, we asked participants to complete individual difference measures for prior knowledge, interest, and attitudes, after which we asked them to complete a multiple text task, involving researching and writing about the topic of overpopulation, using a library of six digital texts. Again, we asked participants to complete tasks involving either formulating an argument or composing a research report. Following the multiple text task, we likewise asked participants to complete the MTSI (Bråten & Strømsø, 2011).

7.2.1. Cued think aloud

In Study 2 we built on findings from Study 1 by adopting a cued think-aloud protocol. The cued think-aloud involved students completing the multiple text task, while hearing a tone in the background, at 60 second intervals. Upon hearing the tone, we instructed students to report anything they were thinking or doing at the time. We chose this methodology as a hybridization of two approaches to capturing strategic processing, commonly used in prior work. One approach, a continuous think-aloud, while capturing online strategy use, may interfere with students’ processing, particularly during complex task completion (List et al., 2016; Cerdán & Vidal-Abarca, 2008). Another approach, retrospective reporting, while allowing students time to reflect on the nature of their strategic engagement, carries the risk of inaccuracy in reporting and positive bias (Van Gog, Paas, Van Merriënboer, & Witte, 2005). We developed the cued think-aloud as a hybrid method capitalizing on the advantages of these two approaches. On the one hand, a cued think-aloud prompted students to report their strategy engagement in a highly contextualized and task-specific fashion, as would a continuous think-aloud. On the other hand, the cued think aloud provided the opportunity for students to reflect and deliberately report their strategic processing, as would a retrospective reporting approach. Moreover, we expected the 60 second interval that we used in the cued think-aloud to result in more minimal interference with processing, as compared to a continuous think-aloud, while offering a somewhat more standardized approach to strategy reporting, allowing students’ utterances to be easily segmented, associated with specific content in texts, and analyzed.

7.2.1.1. Utterances coding. We coded think aloud utterances though a two-step process. First, we coded every utterance for whether or not it was cross-textual in nature. Cross-textual utterances were ones that simultaneously referred to more than one text at a time. We then coded these cross-textual utterances for whether or not they reflected integration. We defined integrative utterances as those that simultaneously drew a connection between two or more texts. For instance, we considered an utterance such as: It has like Empowering Women, More Efficient Food Production, and I’m reading them to find out which one I need to click on, to be a cross-textual utterance that was non-integrative in nature. While this utterance simultaneously considered more than one text (i.e., cross-textual), it did not explicitly draw a connection between them. We contrasted such an utterance with cross-textual utterances that also involved integration. For example: Tying the two articles together in order to display that overconsumption is a grave threat and women are able to control it and have a large factor in it. This utterance not only simultaneously considered two texts, but also identified a complementary relation between them. Two raters coded eight student responses (25.00%), with a Cohen’s kappa of 0.96 for the number of cross-textual and integrative statements identified.

7.2.2. Response coding

As in Study 1, we coded written responses for the type of multiple text model they evidenced. Specifically, we coded responses as demonstrating a mush model, separate representations model, with and without citation, or a documents model approach to conceptualizing multiple texts.

7.2.3. Strategy inventory

Like in Study 1, after they composed a written response, we again asked students to complete the MTSI (Bråten & Strømsø, 2011). Reliability for the five item information accumulation scale was Cronbach’s α = 0.80 and α = 0.87 for the 10 item cross-textual elaboration scale.

8. Results

8.1. Research question 1: Models of multiple texts in written responses

Our first research question examined the types of multiple text models evidenced in students’ written responses. Most commonly, students’ written responses reflected the separate representations model of multiple texts. This was the case for 51.61% of the sample, overall (n = 16), and included both students who had separate representations models with citations (35.48%, n = 11) and without citations (16.13%, n = 5) reflected in their responses. Mush models, reflecting content integration but no citation, occurred less frequently, but nevertheless were reflective of 19.35% of the sample (n = 6). Finally, documents models, reflecting integration, were constructed by 22.58% of the sample (n = 7).

8.2. Research question 2: Integrative think-aloud statements

Our second research question examined the prevalence of integration demonstrated in students’ think-aloud utterances. We examined both the prevalence of cross-textual statements, making simultaneous reference to more than one text, and integrative statements, drawing a connection between texts, reflected in students’ think-aloud utterances. Twenty students (68.97%) included any cross-textual statements in their think-aloud responses, while only eight students (27.59%) produced statements that were integrative in nature. Approximately half of the statements classified as cross-textual in nature (M = 1.55, SD = 1.72), were also considered to demonstrate integration (M = 0.59, SD = 1.18).

8.3. Research question 3: Integrative statements and multiple text models in written responses

In the third research question we examined the association between think-aloud utterances reflecting cross-textual linking or integration and the types of multiple text models that students constructed in their written responses. First, we ran a one-way ANOVA to examine whether the types of multiple text models that students constructed in their written responses was associated with differences in the total number of cross-textual utterances produced during the think-aloud. The overall ANOVA was significant, F(3, 26) = 3.04, p < 0.05, η² = 0.34, indicating a large effect. Using post-hoc analyses using Tukey’s HSD, we found that students constructing documents models produced significantly more total cross-textual utterances, overall, (M = 3.29, SD = 1.98) than students constructing mush models (M = 0.83, SD = 1.33) or separate representations models without citations (M = 0.80, SD = 0.84), ps < 0.05.

Further, we examined whether the types of multiple text model reflected in students’ written responses were associated with differences in the number of integrative utterances reported during the think-aloud. The overall ANOVA was again not significant (p = 0.08).
However, the mean number of integrative utterances reported followed the pattern that we expected, with students constructing documents models (M = 1.57, SD = 1.81) in their written responses also producing the largest number of integrative utterances during the course of the think-aloud, followed by students composing mush models (M = 0.67, SD = 1.21) and models reflecting a separate representation of texts with (M = 0.22, SD = 0.44) and without citations (M = 0.00, SD = 0.00). In Table 8, we summarize the number of integrative and cross-textual statements produced, by multiple text model.

### 8.4. Research question 4. Integrative statements and integrative strategy use

As a final analysis we were interested in examining whether the number of cross-textual and integrative statements that students made during their cued think-aloud were associated with their reports of multiple text strategy use on the MTSTI. Using correlation analyses we found that the total number of cross-textual (p = 0.77) and integrative statements (p = 0.69) that students made during the think-aloud were not significantly associated with students’ reports of information accumulation strategy use on the MTSTI. However, the total number of cross-textual statements that students made during processing was associated with students’ ratings of cross-textual elaborative strategy use, r(27) = 0.38, p < 0.05, indicating a medium to large association. Moreover, there was a significant association between the number of integrative statements that students made during processing and their ratings of cross-textual elaborative strategy use on the MTSTI, r (27) = 0.48, p < 0.05, indicating a medium to large association. Correlation results are presented in Table 9.

### 9. Study 2 discussion

#### 9.1. Research question 1: Models of multiple texts in written responses

In the first research question we were concerned with the type of multiple text models students constructed based on information presented across multiple texts. Our findings from Study 2 echo results produced in Study 1. Specifically, the same multiple text models that we saw produced in Study 1 were also reflected in Study 2. Moreover, the multiple text models constructed as a part of Study 2 were roughly proportionate in prevalence to the models formed as a part of Study 1.

#### 9.2. Research question 2: Integrative think-aloud statements

In our second research question, we considered the types of integration that students demonstrated in their think-aloud utterances. We coded for two types of utterances: utterances that simultaneously referred to multiple texts (i.e., cross-textual utterances) and utterances that drew a relation or a connection between them (i.e., integrative utterances). A critical insight we developed through this coding scheme was the need to distinguish between students considering more than one text during processing and cross textual linking, or integration.

We further found it to be the case that cross-textual and integrative utterances were differentially focused, as a part of multiple text use. Cross-textual utterances were most commonly produced in association with students’ multiple text use behaviors, including text selection and navigation and note-taking. Integrative utterances, as a contrast, tended to be focused on drawing relations between texts’ content and were typically reported during the course of reading, rather than prior to or following text use. Nevertheless, we should note that any cross-textual utterances reported were a minority of all of the behaviors and cognitions that students produced during the course of the think aloud.

#### 9.3. Research quest 3: Integrative statements and multiple text models in written responses

In our third research question we examined the association between integration, as reported during multiple text processing, and integration, as demonstrated in the multiple text models reflected in students’ written responses. We found total number of cross-textual utterances that students reported to differ in association with the type of multiple text model that they constructed. As we expected, documents model construction was associated with the largest number of cross-textual statements produced. However, we found no significant differences across multiple text models in the number of integrative statements uttered during processing. We thought this was likely due to limitations in power, resulting from a small sample size and floor effects associated with the limited degree of integrative statements that students reported during their think-alouds. Indeed, given the 0.25 effect size we found for the differences in the number of integrative utterances students produced across models of multiple texts, post-hoc power analyses found our power (1−β) to only be 0.15. Nevertheless, means followed the patterns we expected. Students composing documents models in their written responses also produced the largest number of integrative statements during their think-alouds. Moreover, aside from documents models, mush model construction was associated with the greatest number of integrative statements uttered during task completion. This pattern of results is consistent with the Documents Model Framework, which conceptualizes both documents models and mush models as reflecting a greater degree of content integration than models reflecting separate representations of multiple texts.

#### 9.4. Research question 4: Integrative statements and integrative strategy use

In Study 2, we used two process measures of integration during
multiple text. These were students’ utterances produced during a cued think aloud and self-reports of strategy use on the MTISI. We found these to be associated with one another. In particular, the number of integrative statements that students produced during their think-aloud was significantly associated with their reports of cross-textual elaborative strategy use, but not with their reports of information accumulation, on the MTISI.

10. General discussion

Looking across both Study 1 and Study 2, we draw three key conclusions. The first point concerns the types of multiple text models that students may construct based on multiple texts. We found evidence for all of the novice models that Britt et al. (1999) suggest emerge when students read multiple texts (i.e., mush model, separate representations model, documents model). This is notable because we are, as far as we know, among the first to holistically consider students’ written responses as reflecting these types of models of multiple texts. Second, we documented a wide range of integration types manifest across students’ written responses. Indeed, in Study 1, the total number of integrative statements that students produced as well as citation use were found to significantly differ across multiple text models; in Study 2, the total number of cross-textual utterances that students reported during task completion were also found to differ in association with the multiple text models that students constructed. Finally, across both Study 1 and Study 2, we found cross-textual elaborative strategy use, as captured via the MTISI, to differ in association with the types of multiple text models reflected in students’ written responses, corresponding to a medium to large effect.

10.1. Models of multiple text

The relative prevalence of each type of multiple text model that students constructed was found to be fairly consistent across Study 1 and Study 2. For instance, we found separate representations models to be most commonly constructed, while mush models were somewhat less prevalent. Indeed, the frequency with which students constructed models reflecting a separate representations view of multiple texts indicates the difficulties that students may experience with integration. We found documents models construction to be evidenced by 34.27% of the sample in Study 1 and 22.58% of the sample in Study 2. This indicates both that expecting undergraduate students to construct documents models of multiple texts is a developmentally appropriate norm and that further instruction and intervention is necessary for a larger proportion of undergraduate students to be able to readily do so. The latter point, associated with the need to improve students’ documents model construction, is notable given the fairly limited degree of integration exhibited even among students composing such models in their written responses.

10.2. Types of integration identified

Our second conclusion addresses the range of integration that students may engage in when learning from multiple texts. Across Study 1 and Study 2, we identified instances of explicit and implicit integration in students’ written responses and found students to draw both complementary and contrasting relations across texts.

10.2.1. Study 1

In Study 1 we taxonomized the level of specificity of integration that may be reflected in students’ responses, with implicit integration more prevalent than explicit integration. This suggests that students may have particular difficulties in explicitly relating texts to one another and identifying the types of relations manifest across texts. Students’ linking of both evidence and themes, or main ideas, across texts was comparatively common, relative to students’ linking of texts according to contextual factors, beyond text content. Missing was students considering author features or points of view in relating texts to one another. This suggests the need to support students to conceptualize and evaluate texts as authored entities, written from some perspective and for some audience (Fox, 2009).

10.2.2. Study 2

In Study 2 we considered whether students were engaging in cross-textual reasoning (i.e., simultaneously considering more than one text) or integration, per se (i.e., determining a specific connection or relation across texts). While a sizable minority of students demonstrated integration in their written responses, as a whole integration was quite limited. Even those students that did include integration in their written responses (i.e., constructing documents models), were only able to do so in a relatively limited number of instances. This points both to the limits in students’ multiple text integration, at present, and to the potential to develop integration by teaching students the range of ways that texts may be connected to one another.

10.3. Indicators of integration

The third conclusion we draw concerns the multidimensional nature of integration. A strength of the analyses we present are that, across two studies, we were able to examine at least three distinct indicators of integration. These included students’ written responses (Study 1 and Study 2), self-reports of strategy engagement (Study 1 and Study 2), and utterances reported during the course of multiple text task completion (Study 2). All three of these measures were associated with one another, to various extents. Specifically, in Study 1 we found the degree of integration included in students’ written responses to be associated with reports of cross-textual elaborative strategy use on the MTISI. In Study 2 we found students reporting of integrative utterances during the course of task completion to be associated with strategy ratings on the MTISI. Jointly, these analyses suggest that integration is a prominent process in students’ multiple text use that dictates both the nature of students’ engagement with multiple texts and manifests in students’ written responses.

10.4. Limitations

Across both Study 1 and Study 2, we must acknowledge a number of limitations. To start, in both studies we examined students’ written responses for evidence of integration. In part, this was an effective method of capturing students’ multiple text model construction. At the same time, written responses may not fully capture the degree of students’ multiple text integration, due to deficits in students’ writing skills (Graham & Perin, 2007; Langer, 2001). Nevertheless, the limitations in integration manifest in students’ written responses were also evidenced in the limited degree of integration demonstrated in students’ utterances during course of the cued think aloud employed in Study 2. These findings suggest that students’ multiple text integration may simply be limited, overall.

Moreover, what constitutes a quality response based on multiple texts demands further clarification. For instance, we considered responses coded into the separate representations, with citation, category to be limited due to their dearth of multiple text integration. At the same time, these responses were organized, included citation, and were constructed based on evidence from texts. In many ways, such responses may be considered to be effective academic products. Consistent with their prevalence, we found such responses to bear a striking resemblance to the five-paragraph essay format, commonly taught in American schools (Campbell & Latimer, 2012; Foley, 1989; Smith, 2006). This suggests the need to further emphasize the importance of integration, as a component of writing based on multiple texts, to students and to teachers alike.

On a related note, we seek to clarify the grain size of what is meant by integration in future work. In this paper, we coded students’ responses holistically as reflecting integration, as well as more
componentially, for the number of integrative statements, discourse connectives, and citations they included. Nevertheless, a lingering question is the extent to which including only a single integrative statement in a response, translates into an overall integrated written composition or a documents model of multiple texts.

In Study 2 we build on findings from Study 1 to confirm that the integration manifest in students’ written responses was, indeed, associated with cross-textual linking during multiple text use, as captured via think-aloud rather than through self-report. To capture students’ integration during multiple text processing, in Study 2 we adopted a cued think-aloud protocol, asking students to report whatever they were thinking or doing at 60 s intervals. While we adopted this procedure to capture students’ processing during the intensive process of reading and writing based on multiple texts, hearing a tone every 60 seconds, may, nevertheless, have proved overly taxing or stressful for students. In part, this may be what is responsible for the somewhat lower degree of documents model construction manifest in students’ written responses in Study 2, and was a unique limitation associated with the second study.

Moreover, a small sample size in Study 2 limited statistical power, and therefore, the results that could be identified. Limited sample size resulted in our identifying a more limited range of connections in students’ written responses and in fewer instances of integration able to be analyzed. Limited sample size also attenuated the statistical associations that we were able to identify and limited the representativeness of our sample and its generalizability. No doubt, we need to replicate Study 2 using a continuous thing-aloud and a larger sample. This remains a critical direction for future work. Indeed, we are currently analyzing data from a larger study (n = 78) that asked participants to type think-aloud comments at designated points throughout multiple text use, rather than cuing reporting at 60 second intervals.

11. Conclusion

In two studies, we use Britt et al. (1999) Documents Model Framework to classify students’ written responses for the type of multiple text model they reflect, among the first studies in the field to do so. Additionally, we introduce a taxonomy of the types of connections that students may form across texts, including evidentiary, thematic, and contextual links, and examine how these may be expressed in students’ writing, using either explicit or implicit devices. To our knowledge, we are among the first studies to examine the nature of multiple text integration at this fine-grained level. Finally, we examine the associations among a variety of measures of integration, including the types of multiple text models featured in students’ written responses, other commonly use integration-related metrics (e.g., citations, discourse connectives), and both inventory- and cued-think-aloud based reports of integrative strategy use. As such, this study represents among the first studies in the field to do so. Are sophisticated students always better? The role of topic-specific personal epistemology in the understanding of knowledge and knowing. Instructional Science, 38, 635–657. https://doi.org/10.1007/s11251-008-9091-4.

Second, we explored the role of topic-specific beliefs about the nature of global warming: Examining the role of topic-specific beliefs about the nature of knowledge and knowing. Instructional Psychology, 6(1), 111–130. https://doi.org/10.1007/s11409-011-9075-7.

We elected to capture students’ interest at the domain level to reflect the nature of overpopulation as an inter-disciplinary topic, arising at the intersection of economics, environmental science, and public policy. At the same time, topic-level measures of interest have commonly been used in prior work and should be considered for inclusion in future iterations of this study (Bråten, McCrudden, Stang Lund, Brante, & Strømsø, 2018; Strømsø et al., 2010). Students rated their interest on a 7-point scale, ranging from not at all interested to very interested. Students’ average interest ratings were 4.03 (SD = 1.43). Cronbach’s alpha reliability for the seven item scale was 0.79.

Attitudes

Students were also asked to report their attitude toward over-population (i.e., the topic of the task). Attitude items reflected either a concern with the threats of overpopulation or the consideration of overpopulation to not be a major threat. Sample items on the attitude scale were: overpopulation is a major problem in the world today and overpopulation is a major threat to political stability. Each item on the attitude scale was endorsed using a 7-point scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Students’ average attitude ratings were 4.71 (SD = 0.91), indicating that students considered overpopulation to be a concern, to a moderate extent. Cronbach’s alpha reliability for the six item scale was 0.72.

First, a hierarchical linear regression was used to predict the total number of integrative statements in students’ written responses. After controlling for task condition in Step 1, students’ total prior knowledge score and average interest and attitude ratings were entered into the model at Step 2. However, this model was not significant (p = 0.17). A second hierarchical linear regression was run to predict the number of discourse connectives included in students’ written responses. Again, task was controlled for in Step 1 and students’ prior knowledge, interest, and attitudes were entered at Step 2. This full model was again not significant (p = 0.35).

A third hierarchical linear regression was conducted to predict the number of citations in students’ written responses. After controlling for task in Step 1, individual difference factors (i.e., prior knowledge, interest, and attitudes) were entered in Step 2. This model was again not significant (p = 0.56).

References


