Understanding students' conceptions of task assignments

Alexandra List[*], Hongcui Du[1], Ying Wang[1]

The Pennsylvania State University, United States

ARTICLE INFO

ABSTRACT

Building on research emphasizing the importance of task in reading performance, the present study examines students' conceptions or definitions of five common academic task assignments (i.e., argument, essay, opinion, summary and research tasks). Results showed that while students have robust schema for some task assignments (e.g., argument, research report), citing a large variety of task elements in defining each of these assignment, other task assignments are only superficially conceptualized by learners (e.g., opinion, essay). While prior work has demonstrated the important role of task assignment in students' text processing and performance, this is among the first studies to associate emergent differences not only with the tasks assigned but also with students' task conceptions. Given the emphasis placed on argument tasks in supporting deeper level processing, we were especially interested in students' conceptions of argument tasks vis-à-vis summary tasks. Students' conceptions of these and other task assignments were compared using Cochran's *Q*, a non-parametric alternative to the repeated measures ANOVA. Finally, linear regression found students' conceptions of an argument task to be associated with performance on an argument writing task assignment. For instance, students considering argument construction to require citation use was associated with their inclusion of citations in their written responses. This study is among the first to examine students' own conceptions of common task assignments and to link task conceptions with performance. Implications for instruction are discussed.

1. Introduction

In a classic psychology study, Pichert and Anderson (1977) instructed students to read the same passage of text about a house from either the perspective of a burglar or from that of a home-buyer. Pichert and Anderson (1977) found that the perspectives that students were assigned were associated with their identification of different information as relevant or not, within the same text, and with perspective-specific differences in recall emerging both immediately after reading and at a one-week post-test. Since this seminal study, researchers have repeatedly documented the role of task assignment or reading goals in text processing and comprehension. Reading goals can be defined as students' assigned (or self-determined) purpose for reading and have been considered to set the criteria according to which reading effectiveness may be judged. Reading goals have been taxonomized as general (e.g., read to study) or specific (e.g., read to define a term; McCrudden & Schraw, 2007) in nature and have been found to be associated with attention allocation and with the identification of relevant vis-a-vis irrelevant information during reading. Moreover, reading goals have been found to support the selection of goal-directed,

active comprehension strategies during reading and to facilitate goal-relevant information recall (Afflerbach & Cho, 2009; McCrudden & Schraw, 2007; Pichert & Anderson, 1977; Pressley & Afflerbach, 1995).

Increasingly, students have been required to comprehend not only single texts, but also to comprehend and integrate multiple textual sources of information to learn about complex topics, like climate change, personal health, or historical controversies (List, Alexander, & Stephens, 2017; Rouet & Britt, 2011). The role of reading goals may be all the more important when students are asked to engage in multiple, rather than single, text use (Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; List & Alexander, 2017; Rouet & Britt, 2011). This is the case for a number of reasons. For one, reading goals direct students' attention to goal-relevant information in text (McCrudden, Magliano, & Schraw, 2011); within a multiple text context, this serves both to reduce the volume of information that students have to manage by allowing irrelevant information to be dismissed and to facilitate the connecting of relevant information across texts. For another, reading goals provide a basis for students' selection, processing, and evaluation of texts (Rouet & Britt, 2011). For instance, relevance to reading goals may guide students' decisions to skim versus read particular sources, or to

integrate information from certain texts, to the exclusion of others. More generally, reading goals help students to determine the standard to which they are required to comprehend and integrate texts and, as a result, dictate students' strategy use and reading behaviors (McCrudden, Magliano, & Schraw, 2010; Schraw, Wade, & Kardash, 1993).

Indeed, the task goals that students are assigned have been found to play a key role in multiple text comprehension, evaluation, and integration (List, Grossnickle, & Alexander, 2016; Stadtler & Bromme, 2007; 2008; Wiley & Voss, 1999). Models of multiple text processing have located task goals, and students' conceptions of those goals, as central to multiple text use (Rouet & Britt, 2011). Most prominently, the Multiple Documents Task-Based Relevance Assessment and Content Extraction (MD-TRACE) model conceptualizes multiple text use as unfolding through a series of five, potentially recursive, steps. Multiple text use begins when students are assigned a particular task. Students then cognitively represent this task, or form a mental model of task demands (i.e., a task model, Step 1). They then determine that they have an information need (Step 2), requiring the accessing of multiples texts. In Step 3 of the MD-TRACE students engage with multiple texts, including (a) determining relevance, (b) processing content, and (c) creating a documents model, or an integrated, cognitive representation of information presented across texts. Students then create a task product (Step 4) and determine the extent to which the task product developed comports with task demands (Step 5).

As specified in the MD-TRACE, students' conceptions of task demands dictate their subsequent multiple text use and response composition. Nevertheless, students' conceptions of task demands, or the task models that students develop, have been under-examined in the literature. This is a concern given that empirical work has found differences in task goals to result in performance differences on a variety of multiple text tasks. For instance, Wiley and Voss (1999) found that students assigned to read multiple texts in order to write an argument, integrated texts to a significantly greater extent than did students asked to write a summary, a narrative essay, or an explanation. Indeed, the facilitative role of certain task goals over others has repeatedly been found in research on multiple text use (Bråten & Strømsø, 2009; Gil, Bråten, Vidal-Abarca, Strømsø, 2010a; 2010b; Le Bigot & Rouet, 2007). Nevertheless, a question remains regarding why certain types of multiple text task assignments improve performance over others. Answering this question requires considering students' own conceptions of different task assignments (i.e., task models) to understand why learners may engage with multiple texts differently when asked to read in order to write an argument vis-à-vis a summary. In this study, we examine students' conceptions of five commonly occurring academic task assignments and associate students' conceptions of one of these tasks (i.e., reading to write an argument) with performance.

1.1. Types of task goals

In their Relevance Taxonomy, McCrudden and Schraw define four types of reading goals that may be assigned to students, forming the basis for learners' relevance determinations (2007). They first distinguish reading goals as specific or general, with *specific* reading goals asking students to identify particular information in texts, and *general* reading goals providing students with a frame of reference through which to approach texts holistically. Specific reading goals include goals asking students to answer the question "what" by locating terms or concepts in text (i.e., targeted segments) as well as goals asking students to answer the question "why" by connecting concepts in text either with other text-based information or with prior knowledge (i.e., elaborative interrogation). General reading goals ask students to either adopt a frame of reference during reading, thereby activating a particular background knowledge schema that would otherwise not be relevant (i.e., perspective), or to read with a focus on learning outcomes, thereby creating the context for reading (i.e., purpose).

All four of these task goals have been examined to varying extents. For instance, students have been assigned to answer high-level versus low-level questions while reading, with the former task assignment found to improve comprehension at post-test (Cerdán, Vidal-Abarca, Martínez, Gilabert, & Gil, 2009). High-level questions were those requiring students to engage in inferencing to connect multiple segments of text to one another, whereas low-level questions had answers explicitly stated in text. According to McCrudden and Schraw's taxonomy, these questions may be said to represent elaborative interrogation vis-à-vis targeted segment specific goals for reading (2007). Elsewhere, Pichert and Anderson (1977) burglar and home-buyer paradigm has been replicated, with burglars and interior decorators' points of view assigned as perspective goals (Kaakinen & Hyönä, 2008). Finally, students have been instructed to read popular science texts either to study or for entertainment, with the studying task resulting in students' engaging in more inferencing and paraphrasing during reading (van den Broek, Lorch, Linderholm, & Gustafson, 2001). Moreover, readers assigned to the studying condition recalled 20% more information than did students reading for entertainment. According to McCrudden and Schraw's framework, these goals represent purpose goals (2007). Such purpose goals, assigning students to read for a general outcome, have been the category of goals most commonly examined in studies of multiple text use.

1.2. Task goals in multiple text use

A variety of task goals have been examined within the context of multiple text use. Typically, these goals have asked students to read in order to produce some kind of written product, consistent with purpose goals identified by McCrudden and Schraw (2007). Building on work by Wiley and Voss (1999), Bråten and Strømsø (2009) likewise examined the role of different task assignments in performance, comparing the efficacy of argument, summary, and global understanding tasks. They found both argument and summary tasks to facilitate multiple text integration, over tasks requiring global understanding. Later, Gil et al. (2010a; 2010b) found that summary tasks better facilitated both multiple text comprehension and integration as compared to argument tasks; a finding that seems inconsistent with other work (Hagen, Braasch, & Bråten, 2014; Le Bigot & Rouet, 2007; Wiley & Voss, 1999). Gil, Bråten, Vidal-Abarca, and Strømsø (2010b) suggest that these findings can be explained in a number of ways, including by task assignment interacting with individual difference factors, like epistemic beliefs; although, limited empirical support was found for this hypothesis. Crucially, Gil et al. (2010b) posit that summary tasks may have facilitated performance over argument tasks because while argument tasks were conceptualized by students as requiring them merely to share their opinions, with limited textual support, summary tasks prompted closer attention to and the more accurate mapping of content in multiple texts (Gil et al., 2010b). Gil, Bråten, Vidal-Abarca, and Strømsø (2010a) suggest that this effect may be especially pronounced for low- rather than high-knowledge learners. This particular supposition is investigated in the present study. Specifically, we are interested in examining students' conceptions of different tasks and, particularly, in understanding the relation between students' conceptions of argument tasks and their performance when asked to read in order to compose an argument. Building on work by Gil et al. (2010a), the relation between argument task conceptions and performance is investigated here in a sample of education majors demonstrating varying levels of knowledge about the topics of the task (i.e., flipped classrooms and academic tracking).

1.3. Present study

The present study contributes to the theoretical literature on multiple text use by explicitly asking students to describe their conceptions of a variety of common academic assignments, with *conceptions* defined

in this study as students' subjective understandings of task demands. While a number of studies have inferred students' task conceptions (e.g., Gil et al., 2010; Wiley & Voss, 1999), this represents the first study, to our knowledge, to explicitly ask students to define how they view a variety of academic assignments. The study had two main parts. First, participants were asked to describe their conceptions of five common academic task assignments, also frequently used in studies of multiple text use. Specifically, participants were asked for their conceptions of argument, summary, opinion, essay, and research tasks. Then, participants were asked to compose an argument based on texts presenting conflicting positions on two controversial education topics (i.e., flipped classrooms and academic tracking). Participants' performance on the argument task was compared to their conceptions of argument tasks.

The five tasks selected were determined based on those commonly used in previous studies of multiple text use. Initially, Wiley and Voss (1999) compared students' performance on argument, summary, explanation, and narrative essay tasks. In this study, we examined many of these same tasks (i.e., argument, summary, narrative essay), with explanation tasks replaced by an assignment asking students to compose a research report. Wiley and Voss (1999) suggest that argument tasks differ from explanation tasks due to their degree of personalization, such that explanation tasks ask students to *tell why this happened* whereas argument tasks ask students to *tell why you think this happened* (*p*. 306). We thought this distinction between personalized and de-personalized assignments was better captured by asking students for their conceptions of writing an argument vis-à-vis a research report (List & Alexander, 2019). Additionally, we asked students for their understandings of what writing an opinion entailed, to contrast this assignment directly with students' conceptions of argument tasks (Gil et al., 2010b; Kobayashi, 2009). Studies of multiple text use have most commonly compared students' performance on summary vis-à-vis argument tasks (Gil et al., 2010a; 2010b; Bråten & Strømsø, 2009; Hagen et al., 2014; Le Bigot & Rouet, 2007), as such students' conceptions of these two task assignments, in particular, were of great interest to us in this study.

An association between students' conceptions of task assignment and performance is expected. For instance, if students believe that argument tasks require them to provide evidence for claims or opinions, we would expect them to compose better justified responses than students considering argument tasks to be synonymous with only providing an opinion. Drawing on prior work examining argumentation, we can define argument quality in a variety of ways. Reznitskaya, Kuo, Glina, and Anderson (2009) developed a rubric to holistically capture argumentation quality. At best, students' arguments should feature key argumentation components (i.e., position, supporting reasons, opposing reasons, elaborations, and rebuttals), consider opposing perspectives, beyond their own, as well as be focused and organized. This rubric has been used to score participants' arguments composed in response to multiple text tasks, with students' performance generally considered to be quite low (Anmarkrud, Bråten, & Strømsø, 2014; Anmarkrud et al., 2013). Elsewhere, arguments have more simply been coded for the extent to which they are dialogic in nature or include evidence to support claims (i.e., arguments), consider counter-arguments or opposing points of view, and offer refutations (Iordanou and Constantinou, 2014; Nussbaum & Schraw, 2007). Given the limitations in students' argumentation quality previously identified (Anmarkrud et al., 2014; Anmarkrud et al., 2013), this latter, more simplified coding scheme is used to score students' argument construction in this study. Argument quality is compared to students' conceptions of what argument tasks entail.

We elected to focus on validating students' performance on argument tasks, relative to their conceptions of such tasks, for three primary reasons. First, argument tasks have been the focus of considerable research and broad curricular efforts (Kuhn, 2010; Kuhn & Moore, 2015; Newell, Beach, Smith, & VanDerHeide, 2011; Wolfe, 2011). Second,

argument tasks have been found to foster improved multiple text comprehension and integration, in particular, as compared to a variety of other task assignments (e.g., summary tasks, Bråten & Stromso, 2009; Gil et al., 2010a; 2010b; Wiley & Voss, 1999). Finally, and perhaps most importantly, across studies examining argumentation, generally, and argumentative writing based on multiple texts, more specifically, students have been found to perform poorly (e.g., Monte-Sano & De La Paz, 2012; Nussbaum & Schraw, 2007), including frequently composing only one-sided arguments or arguments evidencing a "my-side" bias (e.g., Wolfe & Britt, 2008; Wolfe, Britt, & Butler, 2009); this lack of two-sided argumentative writing has often times been attributed to students' lack of knowledge regarding what argumentative tasks entail and difficulties separating the demands of writing an argument from those of composing an opinion-based or persuasive essay (Wolfe et al., 2009). Given that argument tasks have commonly been used in prior work and, students' poor performance on these attributed to a lack of task-specific understandings of what argumentation requires, in this study we were specifically interested in examining the extent to which students with different connections of argument tasks differentially performed when asked to compose an argument based on multiple texts.

We have the following research questions:

1. What are students' conceptions of common academic task assignments?
2. What are the similarities and differences in how students conceptualize common academic task assignments?
3. What is the association between students' conceptions of argument tasks and performance?

2. Method

2.1. Participants

Participants were 132 undergraduate students enrolled in a Mid-Western University in the United States, with an average age of 21.16 years (*SD* = 4.24). In terms of gender, 17.74% (*n* = 22) of participants were male, 81.45% (*n* = 101) were female, and 0.81% (*n* = 1) of subjects reported their gender as other. Among students reporting race/ethnicity, 87.50% (*n* = 105) of the participants were White, 10.00% (*n* = 12) of them were African American, and 2.50% (*n* = 3) of reported Hispanic/Latino ethnicity. Participants represented a variety of class standings; 13.71% (*n* = 17) of participants were freshmen, 36.29% (*n* = 45) were sophomores, 24.19% (*n* = 30) were juniors, 25.00% (*n* = 31) were seniors, and 0.81% (*n* = 1) were advanced special students, taking courses beyond their senior year. The majority of participants were pursuing education or education-related majors, and all were enrolled in educational psychology courses in the University. Twelve students did not provide demographic information. These data are taken from a larger study examining the role of attitudes in text processing.

2.2. Measures

This study had two main parts. First, participants completed an assignment analysis task, asking them to describe what they believed common academic task assignments required them to do. Then, participants completed an argument task twice, addressing two different topics in education (i.e., academic tracking, flipped classrooms). Each argument task included three components. First, participants reported their attitudes on the topic of the task, with attitudes defined as individuals' bi-polar evaluations of an attitude-object, in this case various policies in education (Ajzen & Fishbein, 1977). Then, participants read two texts providing directly opposing views on the topic. After reading the conflicting texts, participants were again asked to report their attitudes on the topic (e.g., supporting or opposing the introduction of flipped classrooms), as well as to provide a justification for their

response.

2.2.1. Prior knowledge

Prior knowledge measures were used to gauge participants' general familiarity with the topics of the two argument tasks (i.e., flipped classrooms, academic tracking). Specifically, participants were asked to respond to the prompt: *Please tell us everything you know about flipped classrooms/tracking in schools. If you're not familiar with flipped classrooms/tacking, please write 'I don't know'.* Students' responses to each question were scored as correct or incorrect, according to whether or not each concept was appropriately described, with accurate and relevant information provided. For instance, responses describing tracking as "separating students based on ability/achievement" or as "it sets students up for vocation, trade school, or college," were both scored as correct. A binary coding scheme was used because we expected participants to have some general familiarity with each of these topics, but to lack deep-level disciplinary or schematic knowledge; therefore, only general familiarity was used as a measure of prior knowledge. Two raters coded 30 student responses (22.73%) for each text topic. Cohen's kappa inter-rater reliability was .92 for the flipped classrooms topic and .80 for the academic tracking topic.

2.2.2. Assignment analysis task

Participants were first asked to describe what they were expected to do to complete five different academic writing assignments, typical of those assigned in high school and college. Specifically, students were asked to define what it meant to write an argument, essay, summary, opinion, and research report. Students were first presented with an assignment statement like, *Please write an argument about a topic in child development (*e.g.*, nature* vs. *nurture; Piaget; Vygotsky)*. Then, participants were asked: *What do you think this assignment is asking you to do?* Participants were further instructed to explain what each assignment was asking them to do, as explicitly as possible. Participants were asked to define all five assignment types; writing assignments were presented in random order.

2.2.3. Coding

Students' responses were coded using a semantic coding scheme, with categories developed based on the language or terms students explicitly used in their task conceptions. The coding scheme was developed through a three-step process. First, three researchers read through all of the participant responses independently, segmented each response into idea units, and assigned a category to each idea unit. At this stage, categories were developed to reflect, as closely as possible, the specific terms participants used in their conceptions of task assignments. Researchers developed these initial categories independently of one another, but were familiar with the field of goal oriented reading and multiple text comprehension. Second, the three researchers compared the categories they developed independently, to create a common set of coding categories. Through this process, similar categories were combined and categories with low response frequencies were deleted, recoded, or merged into larger categories. The three researchers reviewed all student responses for adherence to this updated set of categories. At this stage, initial inter-rater reliability was computed to determine a sufficient level of agreement. Finally, a master coding scheme, reflecting a total of 24 categories, was developed. Raters again reviewed all student responses to finalize their categorization and final inter-rater agreement was computed. Based on all student responses, exact agreement between raters ranged from 92.65% to 93.02% for the argument task (Cohen's $\kappa > .85$); 93.74% to 95.14% for the essay task (Cohen's $\kappa > .87$); 94.48% to 95.14% agreement for the summary task (Cohen's $\kappa > .89$); 92.14% to 95.60% agreement for the opinion task (Cohen's $\kappa > .84$), and 93.45% to 95.03% for the research task (Cohen's $\kappa > .87$). All disagreements were resolved through discussion. This final set of 24 coding categories is defined in Table 1. These categories corresponded to the task elements that

students reported in defining each academic writing assignment.

2.2.4. Attitudes

Participants were asked to report their attitudes and attitude certainty in association with each of the two topics of the argument task (i.e., flipped classrooms, academic tracking). Because participants may have been unfamiliar with either of these topics, each topic was first briefly defined. For instance, participants were told:

> *"Some people think that flipped classrooms are effective in improving student learning. When classrooms are flipped, students learn new content at home, independently, and use class time for guided practice. For example, students may view a lecture on new math content at home and use class time to practice new skills learned, where before practice would have been assigned as homework."*

Students were then asked: *do you think that flipped classrooms/academic tracking improve student learning*, in response to which they selected *yes, no,* or *I'm not sure*. Then, participants were asked to report their attitude certainty using a three item scale. Both prior to and following reading students were fairly evenly distributed in their attitudes supporting or opposing flipped classrooms and academic tracking. See Table 2 for an attitude summary. Average certainty scores, associated with each of the two topics, were used as control variables in regression models for Research Question 3.

2.2.5. Argument task

Participants completed two argument tasks, one on the topic of flipped classroom and the other on academic tracking. The order of the two argument tasks was counterbalanced across participants. The argument task involved participants reading two conflicting texts on each of the target topics and then writing an argument to justify their position on each controversial topic.

2.2.5.1. Conflicting texts. Students were asked to read two conflicting texts about each topic (i.e., flipped classroom and academic tracking). The texts were structured to take an explicit stance in favor of or opposing each target topic and to provide four reasons supporting their position. Each set of conflicting texts were isomorphic and, while directly conflicting, did not explicitly reference one another. Texts were attributed to trustworthy sources and presented in counterbalanced order. Students were randomly assigned to view texts as attributed either to professors of education or to teachers, both of whom would have knowledge about target issues. Although examining differences in text attribution is beyond the scope of this paper, text condition is controlled for in analyses. A sample text is included in Appendix A.

Texts were matched on structure and length, and reflected a reading difficulty level appropriate for an undergraduate audience. Conflicting texts on flipped classrooms were 282 and 284 words in length and had a Flesch-Kincaid Grade Level of 12.4 and 13.3. The two conflicting texts on academic tracking were 264 and 279 words in length and had a Flesch-Kincaid Grade Level of 12.3 and 12.5.

2.2.5.2. Attitude justification. After students read each set of conflicting texts, participants were asked to respond to prompts such as: *You're asked to help your principal decide whether or not to implement academic tracking at the high school level. Should your principal implement academic tracking*; and to endorse, *Yes, No,* or *I'm not sure* in response. Participants were then asked to justify their chosen stance by writing an argument about whether or not their principal should implement flipped classroom or academic tracking at their school. Participants were further instructed to include specific arguments from the texts provided, including information that both supported and conflicted with their chosen position. Participants' justifications or arguments composed in response to conflicting information are the primary outcome of interest in this study.

Table 1
Coding Scheme for assignment definitions.

Category	Description
Recursive	Definitions using same terms as task assignment
Writing	Task elements associated with length, formatting, or style
Research	Task elements associated with conducting research or gathering information
Source use	Task elements associated with the use of textbooks, research articles, or other sources
Citation	Task elements associated with reference use, citation, or quotation
Elaboration	Task elements associated with explaining a topic in depth or in great detail
Opinion	Task elements associated with forming a personal opinion or expressing personal beliefs or views
Justification	Task elements associated with providing facts or evidence in support of claims
Summary	Task elements associated with briefly describing or overviewing a topic
Main idea	Task elements associated with identifying the main idea or important information
Reword	Task elements associated with rewording or rephrasing information or expressing information in one's own words
Demonstrating knowledge	Task elements associated with demonstrating learning or knowledge
Choosing a topic	Task elements associated with selecting a topic of focus
Taking a stance	Task elements associated with adoption a position or taking a stance on a controversy; developing a thesis
Argumentation	Task elements associated with making argumentation moves, including persuasion, defense, or refutation
Content	Task elements associated with providing specific information in writing (e.g., definitions, examples)
Informing audience	Task elements associated with a concern for audience or proving information to readers
Information transformation	Task elements associated with thinking critically about, analyzing, applying, or synthesizing information
Personal experience	Task elements associated with writing about personal experiences
Both sides	Task elements associated with considering both sides of controversies, comparing and contrasting, or considering the benefits and drawbacks of topics
Subjective	Task elements associated with an emphasis on subjectivity or personal perspective and the exclusion of objective facts
Objective	Task elements associated with an emphasis on objectivity and the exclusion of opinions or bias
No addition	Task elements associated with not providing excessive detail or information; limited elaboration
Other	Task elements not otherwise coded

Table 2
Summary for students' attitude.

Topic	Yes		No		I'm not sure		Total
	N	Percent	N	Percent	N	Percent	
Flipped classroom	44	33.33%	54	40.91%	33	25.00%	132
Academic tracking	59	44.70%	43	32.58%	27	20.45%	132

2.2.5.3. Coding. A coding scheme was developed to score students' responses to the argument task. Responses were scored on a five-point scale, shown in Table 3. The coding scheme reflected the volume of evidence participants provided in supporting their own positions as well as the extent to which they considered and refuted conflicting points of view. Responses coded as *non-responsive* reflected participants providing unjustified claims or irrelevant information. The *limited* category represented responses including only a single justification for students' claims. *Elaborated* responses provided multiple reasons or pieces of evidence to support claims; however, these responses were only one-sided, not considering conflicting evidence. Responses scored as *multiplist* provided both multiple justifications for claims and considered conflicting perspectives. However, students with multiplist responses did not refute or otherwise evaluate any of the conflicting arguments provided. Finally, *dialogic* responses included multiple

justifications for claims, the consideration of counter-evidence, and its refutation or evaluation. Students with dialogic responses not only considered both sides of controversial topics but also demonstrated the critical analysis or evaluation of evidence.

Based on the instructions that students received and definitions of argument quality (Reznitskaya et al., 2009), we were particularly interested in the extent to which students included text-based evidence in their written responses. The inclusion of text-based evidence could be differentiated from written responses only based on anecdotal experience (e.g., being a student in a flipped classroom) or world knowledge. Students' responses received a binary score of 0 or 1 according to whether or not they included text-based evidence in their responses. Responses were further coded for their inclusion of citations, or direct attributions of information to sources of origin (e.g., referencing text authors by name or title). Due to the low level of citation use in students' responses, overall, citations were binary scored as present or not.

Two raters scored all students' responses. Inter-rater agreement for the flipped classroom topic was 93.94% for overall argumentation scores and 98.48% for the number of citations included. For the academic tracking topic, inter-rater agreement was 88.64% for overall argumentation scores and 99.24% for citation inclusion. Any discrepancies between coders was resolved through discussion.

Table 3
Coding scheme for argument responses.

Category	Score	Description	Sample response
Non-responsive	0	Non-responsive/Unjustified claim	"I think that it is a good idea to use."
Limited	1	Single justification provided	"Tracking enables students to get the most out of where they are at educationally."
Elaborated	2	Multiple justifications provided	"The principal should implement a flipped classroom because research has shown that flipped classrooms help the student by giving them more responsibility and they help the teachers because it takes off some of the work load."
Multiplist	3	Multiple justifications provided; counter-arguments presented	"Academic tracking has both positives and negatives. On the plus side, it lets more different students learn at the pace and level they need to. On the down side, it does nothing for a student's self-esteem and they will think they don't have to try."
Dialogic	4	Multiple justifications provided; counter-arguments considered and refuted	"While flipped classrooms may benefit the academically advantaged and advanced, they put the disadvantaged at a much proportionally greater disadvantage. It puts a greater burden on the student to learn at home, which is unlikely to happen for underperforming students."

2.3. Procedures

First, participants were asked to complete the assignment analysis task (i.e., to define each of five commonly occurring academic task assignments) and measures of prior knowledge and attitudes. Then, participants were asked to complete the argument task twice, corresponding to two different controversial topics (i.e., flipped classroom and academic tracking). Topics were presented in counterbalanced order. For each of the argument tasks, participants read two conflicting texts on each topic and then wrote an argument justifying their chosen point of view. Participants did not have texts available when composing their written responses. All study materials were presented online using Qualtrics software (Qualtrics, Provo, UT). Participants completed the study at a time and location of their choosing. The study took approximately 30–45 min to complete.

2.4. Analyses

Three sets of analyses were carried out as a part of this study. To do this, the task elements that students introduced in defining each task assignment were binary scored as present or absent. These binary scores represented whether or not students considered particular task elements (e.g., citation) to be necessary for completing a target task assignment (e.g., summary task, research report). The first research question descriptively examined the task elements that students considered to be key to completing each of the five task assignments examined. The second research question examined the extent to which different task elements were proportionately cited in defining each of the five target task assignments. Cochran's Q, a non-parametric alternative to the repeated-measures ANOVA for binary data, was used to test the proportionality of task elements cited across task assignments, with McNemar's Test used for post-hoc pair-wise comparisons. Finally, the third research question examined students' conceptions of argument tasks, in particular, and associated these with performance on two argument composition tasks. Three metrics of task performance were examined: overall argument quality as well as dichotomous outcomes reflecting whether or not students included citations and text-based evidence in their written responses. Three sets of models were run. First, in modeling overall argument quality, step-wise multiple regressions were run in association with each topic. Prior knowledge, attitude certainty, and text condition were controlled for in Step 1 and binary scores for students' citation of task elements associated with effective argumentation (i.e., taking a stance, providing justification, engaging in argumentation, source use, and considering both sides of an issue) were entered at Step 2.

Second, step-wise binary logistic regressions were run to predict whether or not students included citations in their written responses, for each topic. Prior knowledge, attitude certainty, and text condition were controlled for in Step 1 and binary scores for students defining argument tasks as requiring citation or not were entered in Step 2. Third, step-wise binary logistic regressions, one for each topic, were run to predict whether or not students included text-based evidence in their written responses. Again, prior knowledge, attitude certainty, and text condition were controlled for in Step 1 and binary scores for students' argument conceptions as requiring research and source use were entered at Step 2. Analyses run for Research Question 3 are summarized in Table 4.

3. Results

3.1. Research question 1: Students' conceptions of common academic task assignments

The first research question examined students' conceptions of five common academic task assignments. Specifically, students were asked to define what writing an argument, essay, opinion, summary, or

Table 4
Summary of analyses for Research Question 3.

	Flipped Classrooms	Academic Tracking
Multiple linear regression to predict overall argument quality		
Step 1: Controls		
	Prior knowledge[b]	Prior knowledge[b]
	Attitude certainty	Attitude certainty
	Text condition[b]	Text condition[b]
Step 2: Argument Task Elements		
	Taking a stance[b]	Taking a stance[b]
	Providing justification[b]	Providing justification[b]
	Argumentation[b]	Argumentation[b]
	Source use[b]	Source use[b]
	Considering both sides[b]	Considering both sides[b]
Binary logistic regression to predict citation		
Step 1: Controls		
	Prior knowledge[b]	Prior knowledge[b]
	Attitude certainty	Attitude certainty
	Text condition[b]	Text condition[b]
Step 2: Argument Task Elements		
	Citation[b]	Citation
Binary logistic regression to predict inclusion of text-based evidence		
Step 1: Controls		
	Prior knowledge[b]	Prior knowledge[b]
	Attitude certainty	Attitude certainty
	Text condition[b]	Text condition[b]
Step 2: Argument Task Elements		
	Research[b]	Research[b]
	Source Use[b]	Source Use[b]

Note: [b]Denotes a binary variable.

research report entailed.

3.1.1. Argument

When asked to define an argument task, four types of task elements were commonly cited. In particular, 40.15% ($n = 53$) of students reported the need to *take a stance* when writing an argument. This category corresponded to students reporting the need to take a position on a controversial topic or to decide whether one side or another of a topic was correct or incorrect. Students also frequently reported providing *justification* as a necessary component of writing an argument. The need for justification was cited by 34.09% ($n = 45$) of respondents and reflected students reporting the need to support arguments with evidence, facts, or some other form of substantiation. Additionally, 27.27% ($n = 36$) of students stated that an argument should include *both sides* of a controversial topic. This was reflected in participants describing the need to consider conflicting positions or to evaluate the benefits and drawbacks of a particular topic. Finally, 25.00% ($n = 33$) of the students commonly considered writing an argument to involve *argumentation*. For example, one student reported that writing an argument involved: *making it persuasive enough that you could convince someone to side with you*. Other argumentation components cited included defending a position or rebutting counterarguments. All of the task elements that students reported when asked to describe what writing an argument entailed are summarized in Table 5.

3.1.2. Essay

Students most frequently reported two types of task elements when defining what writing an essay entailed. First, 42.42% ($n = 56$) of students considered writing an essay to involve *writing*. This category included concerns regarding writing format and length. Further, 28.79% ($n = 38$) of students reported *elaboration* as a necessary component of writing an essay. This category included students reporting the need to explain or write in-depth about a particular topic. All of the task elements that students reported when asked to describe what writing an essay entailed are summarized in Table 6.

Table 5
Argument task definitions.

Category	Sample Response (Argument)	N (%)
Recursive	"This is asking me to argue"	18 (13.64%)
Writing	"I feel that this is asking me to write a paper on this topic"	7 (5.30%)
Research	"In this assignment, the teacher is most likely wanting me to do a research paper"	8 (6.06%)
Source Use	"I would just take information from the book"	14 (10.61%)
Citation	"I would need a certain amount of citations"	2 (1.52%)
Elaboration	"elucidate upon a well thought out and constructed (rhetorical) question"	9 (6.82%)
Opinion	"This assignment is asking me to give my own opinion on one of this topics"	22 (16.67%)
Justification	"provide evidence and support about why that is the correct view"	45 (34.09%)
Summary	"summarize what the argument is"	1 (0.76%)
Demonstrating Knowledge	"show that we fully understand the concepts"	3 (2.27%)
Choosing a Topic	"pick any topic about child development"	20 (15.15%)
Taking a Stance	"This assignment is asking me to take a stance on a child development related topic"	53 (40.15%)
Argumentation	"Persuade people why the view is correct or wrong."	33 (25.00%)
Content	"I should probably have at least 3 specific supporting points"	5 (3.79%)
Informing Audience	"I would explain. in great depth so the reader has background knowledge on the topic"	2 (1.52%)
Information Transformation	"Actually think about how it applies to real people and real situations"	6 (4.55%)
Both Sides	"It includes research for both sides"	36 (27.27%)
Subjective	"An argument is a topic that may have multiple opinions."	1 (0.76%)
Objective	"prevent the action of bias."	3 (2.27%)
Other	"people believe you can have one and not the other but in reality, you need both you fully develop your child"	27 (20.45%)

3.1.3. Opinion

Two task elements were commonly reported when students were asked to define what writing an opinion required. First, 56.06% ($n = 74$) of the students reported writing an opinion to involve forming an *opinion*. As a note, students who explicitly wrote that opinion writing involved forming an opinion had responses coded into the recursive category. Rather, students whose responses were coded into the opinion category defined writing an opinion as involving the sharing of personal thoughts, beliefs, or views on a certain topic. This category included students' responses like: *write about my views and thoughts about a topic* or *write what I personally think*.

Nevertheless, students' conceptions of opinion tasks were most often coded into the *recursive* category. This category included 45.45% ($n = 60$) of students, who described an opinion assignment using the same language as was used in the assignment instructions. Such responses included: "to write an opinion about a topic in child development." All of the elements reported by students when defining this task can be seen in Table 7.

3.1.4. Research report

In defining what a research report entailed, students most commonly reported three types of task elements. First, 38.64% ($n = 51$) of students reported task elements that were coded as *recursive* in nature. These corresponded to students defining writing a research report as involving writing about research. Students also frequently considered writing a research report to entail *source use*. This task element was reported by 38.64% ($n = 51$) of the sample and included statements like: *use outside articles, normally peer reviewed, to come to a conclusion based on my topic.* More generally, this category reflected students expressing a need to use information sources, of various kinds, in composing a response. Additionally, 39.39% ($n = 52$) of students considered composing a research report to involve *writing*, reflecting a concern with format or length. All of the task elements that students reported when asked to describe what a research report entailed are summarized in Table 8.

3.1.5. Summary

When asked to define a summary assignment, four types of task

Table 6
Essay task definitions.

Category	Sample Response (Essay)	N (%)
Recursive	"Write an essay on child development"	18 (13.64%)
Writing	"An essay would be a 3 page paper or more on the topic of choice"	56 (42.42%)
Research	"find more information outside of the textbook"	22 (16.67%)
Source Use	"This includes providing information from the textbook or official websites"	24 (18.18%)
Citation	"I'm not sure if citations are required"	1 (0.76%)
Elaboration	"more in-depth detail"	38 (28.79%)
Opinion	"most likely consisting of my thoughts"	19 (14.39%)
Justification	"with main facts backing it up"	22 (16.67%)
Summary	"This essay is asking me to give an overview of this particular topic"	11 (8.33%)
Main Idea	"discussing their key beliefs"	2 (1.52%)
Demonstrating Knowledge	"Discuss your knowledge on the topic"	13 (9.85%)
Choosing a Topic	"talking about a chosen topic in child development"	18 (13.64%)
Taking a Stance	"I am being ask to write what is best, nature (biological, natural) or nurturing (emotional, thoughtful)"	3 (2.27%)
Argumentation	"attempt to persuade them why you are right and they are wrong or vice versa"	4 (3.03%)
Content	"construct a story almost to explain nature vs. nurture"	21 (15.91%)
Informing Audience	"focused on educating the reader about said topic or an element of the topic"	6 (4.55%)
Information Transformation	"how it connects to child development"	11 (8.33%)
Personal Experience	"it is an assignment that could be asking for your own personal experiences"	4 (3.03%)
Both Sides	"I wouldn't have to side with any of them"	6 (4.55%)
Subjective	"however no extensive research is necessary"	5 (3.79%)
Objective	"I should leave any opinions"	8 (6.06%)
Other	"does not give you the intention of the essay"	28 (21.21%)

Table 7
Opinion task definitions.

Category	Sample Response (Opinion)	N (%)
Recursive	"They want us to give our opinion of a certain topic in child development using past knowledge or experiences."	60 (45.45%)
Writing	"It should probably be about a paragraph."	12 (9.09%)
Research	"I would need to do some research and include it in my paper."	5 (3.79%)
Source Use	"I should only use the book to back up my opinion."	7 (5.30%)
Elaboration	"This assignment is asking the student to choose an area in childhood development, explain it"	3 (2.27%)
Opinion	"Discuss what your views are on a topic in child development."	74 (56.06%)
Justification	"To write about your thoughts on a topic and explain why."	18 (13.64%)
Summary	"Based on research, I would summarize the topic and give my opinion."	1 (0.76%)
Reword	"This assignment is asking me to explain in my own words a topic in child development."	1 (0.76%)
Demonstrating Knowledge	"They want us to give our opinion of a certain topic in child development using past knowledge or experiences."	12 (9.09%)
Choosing a Topic	"I am expected to pick a topic in child"	12 (9.09%)
Taking a Stance	"It is asking me to state whether i agree or disagree and who's side i am on regarding the subject."	24 (18.18%)
Argumentation	"Write a persuasive essay about my opinions about using support, either from our textbook or research I complete."	3 (2.27%)
Content	"I would define the topic, then give my opinion using support from the theory to justify my opinion."	4 (3.03%)
Information Transformation	"This assignment is asking me to use all of the information I learned about the topic to come to my own conclusion about it and state what I think."	8 (6.06%)
Personal Experience	"Write a paper about child development using your opinions, beliefs, values, and experiences."	5 (3.79%)
Both Sides	"maybe there are some parts I agree on and others I disagree on."	2 (1.52%)
Subjective	"This is just asking your opinion on a particular topic."	25 (18.94%)
Other	"One parent is better than two parents or Los Angeles is worse to raise children than Muncie Indiana."	26 (19.70%)

elements were most commonly reported. First, students defined writing a summary *recursively*. This was reported by 34.85% (*n* = 46) of students, who defined the summary assignment using the same terms as were provided in the assignment instructions. The second most commonly reported task element corresponded to the *summary* category, cited by 34.09% (*n* = 45) of students. Instead of defining the assignment only as *summarizing the topic*, responses coded into the summary category included descriptions such as providing a *short overview,* a *brief discussion*, or *the basics of the topic*. The third task element most commonly cited was *writing*, reported by 29.55% (*n* = 39) of students. Finally, 18.18% (*n* = 24) of students reported that a summary assignment involved identifying *main idea(s)*. Responses in this category included the need to give: *key points* or *important information* when writing a summary. Table 9 shows all of the task elements reported by students in association with this assignment.

3.2. Research question 2: Task elements across task assignments

The second research question examined the consistency of task elements identified across the five different task assignments. Only task elements reported by at least 10% of the sample, overall, were examined across task conditions. Specifically, the prevalence of nine task

elements was examined (i.e., recursive definition, writing, research, source use, elaboration, opinion, justification, choosing a topic, and taking a stance). Cochran's Q, a nonparametric alternative to the repeated measures ANOVA, was used to determine the extent to which various task elements were disproportionately reported across various task assignments. McNemar's tests, with alpha adjusted for multiple comparisons, were used to perform planed, post-hoc, pair-wise comparisons. Students' reporting of various task elements across the five different task assignments are summarized in Table 10.

3.2.1. Recursive definition

Cochran's Q determined that the proportion of students defining the various task assignments recursively varied significantly across tasks, Cochran's $Q(4) = 69.83$, $p < .001$. McNemar's Test for paired samples was used for post-hoc analyses, with alpha adjusted to $\alpha < 0.005$ for multiple comparisons. Results determined that students defined tasks recursively in the case of the opinion task assignment more frequently than they did in the case of the argument $[\chi^2(132) = 35.02, p < .001]$ and essay $[\chi^2(132) = 32.33, p < .001]$ task assignments. Moreover, summary tasks were defined recursively more frequently than were argument $[\chi^2(132) = 17.36, p < .001]$ and essay $[\chi^2(132) = 18.23, p < .001]$ tasks. Finally, research task assignments were also defined

Table 8
Research report task definitions.

Category	Sample Response (Research)	N (%)
Recursive	"I feel that this assignment is asking me to do research"	51 (38.64%)
Writing	"The research paper would be more lengthy than the others"	52 (39.39%)
Research	"I think this assignment is asking you to collect data through scholarly journals"	36 (27.27%)
Source Use	"use scholarly sources to inform"	51 (38.64%)
Citation	"with actual cited sources as evidence"	23 (17.42%)
Elaboration	"in depth discussion about an aspect of child development"	20 (15.15%)
Opinion	"give my opinion on it"	2 (1.52%)
Justification	"include facts about a certain topic"	27 (20.45%)
Summary	"summarize that research"	4 (3.03%)
Demonstrating Knowledge	"then use my new-found knowledge"	7 (5.30%)
Choosing a Topic	"This assignment is asking you to choose a topic in Child Development"	16 (12.12%)
Taking a Stance	"finally conclude the paper with whether or not they were correct on the topic in question"	7 (5.30%)
Argumentation	"This assignment is asking you to create a thesis"	5 (3.79%)
Content	"Write a paper with facts about the field and the different theories"	8 (6.06%)
Informing Audience	"then write about what I find"	23 (17.42%)
Information Transformation	"This assignment is asking me to apply a theory in child development to an experiment"	24 (18.18%)
Both Sides	"compare and contrast the two"	3 (2.27%)
Objective	"that is not your own opinion"	12 (9.09%)
Other	"nature is important but not as much as nurture is"	12 (9.09%)

Table 9
Summary task definitions.

Category	Sample Response (Summary)	N (%)
Recursive	"They just want us to briefly summarize a given topic of child development."	46 (34.85%)
Writing	"giving the key points in a paragraph or so."	39 (29.55%)
Research	"Take the information from class or a student's own research and pick out the most important concepts from it."	5 (3.79%)
Source Use	"Take information presented in a textbook/article/other source."	22 (16.67%)
Elaboration	"To explain a topic in your own words."	20 (15.15%)
Opinion	"maybe give my opinion."	3 (2.27%)
Justification	"This is just asking facts about a topic to be shared in a paper form."	10 (7.58%)
Summary	"briefly talk about a specific subtopic about child development."	45 (34.09%)
Main Idea	"It wants me to take the main points of the"	24 (18.18%)
Reword	"in your own words, hit the main points of it."	17 (12.88%)
Demonstrating Knowledge	"summarize what you have learned and studied in class."	16 (12.12%)
Choosing a Topic	"Pick a topic and outline the main points of it."	17 (12.88%)
Content	"To briefly explain and define a topic in child development."	13 (9.85%)
Informing Audience	"let readers know what it means and what it is about."	3 (2.27%)
Information Transformation	"how it is applied (how you would model the topic in your own classroom) in your own words."	8 (6.06%)
Both Sides	"This would be just explaining the different sides."	3 (2.27%)
Objective	"Based upon the study itself; not my view of it."	2 (1.52%)
No Additions	"Write a small paragraph that briefly explains a topic without going into too much depth."	11 (8.33%)
Other	"Titanic for example, is a very long movie."	13 (9.85%)

recursively more commonly than were argument [χ^2(132) = 21.79, $p < .001$] and essay [χ^2(132) = 22.76, $p < .001$] tasks. In particular, when asked to write opinions, research reports, and summaries, between 34.85% ($n = 46$) and 45.45% ($n = 60$) of students defined these recursively, as compared to 13.64% ($n = 18$) of students defining both the argument and the essay tasks in synonymous ways.

3.2.2. Writing

The prevalence of students considering writing to be a necessary part of task completion was examined across the five different task assignments. Cochran's Q indicated that there were significant differences across the five different task assignments in the proportion of students reporting these to require a particular length or format, Cochran's $Q(4) = 92.02$, $p < .001$. Post-hoc analyses using McNemar's Test determined that students associated writing significantly more commonly with the essay task, than they did with the argument [χ^2(132) = 45.18, $p < .001$] and opinion [χ^2(132) = 38.52, $p < .001$] tasks. Additionally, students more commonly associated writing with the summary task than they did with the argument

[χ^2(132) = 26.69, $p < .001$] and opinion [χ^2(132) = 17.33, $p < .001$] tasks. McNemar's Test also determined that students considered writing to be more pertinent to the research task assignment than the argument [χ^2(132) = 35.20, $p < .001$] and opinion [χ^2(132) = 29.25, $p < .001$] task assignments. Specifically, 42.42% of students ($n = 56$) reported writing to be a necessary task element for completing the essay assignment, as compared to 39.40% of students ($n = 52$) for the research task, 29.55% of students ($n = 39$) for the summary task, 9.10% ($n = 12$) of students for the opinion task, and 5.30% ($n = 7$) of students for the argument task.

3.2.3. Research

Cochran's Q further determined that the need to conduct *research* was a task element disproportionately reported across the five different task assignments, Cochran's $Q(4) = 56.40$, $p < .001$. Post-hoc analyses using McNemar's Test indicated that students more frequently reported conducting research to be necessary for writing a research report than for composing an argument, [χ^2(132) = 21.44, $p < .001$], an opinion [χ^2(132) = 24.32, $p < .001$], or a summary [χ^2(132) = 21.95,

Table 10
Overall table of task elements reported across task assignments.

	Argument	Essay	Opinion	Summary	Research	Total
Recursive	18 (13.64%)	18 (13.64%)	60 (45.45%)	46 (34.85%)	51 (38.64%)	193 (29.24%)
Writing	7 (5.30%)	56 (42.42%)	12 (9.09%)	39 (29.55%)	52 (39.39%)	166 (25.15%)
Research	8 (6.06%)	22 (16.67%)	5 (3.79%)	5 (3.79%)	36 (27.27%)	76 (11.52%)
Source use	14 (10.61%)	24 (18.18%)	7 (5.30%)	22 (16.67%)	51 (38.64%)	118 (17.88%)
Citation	2 (1.52%)	1 (0.76%)	0	0	23 (17.42%)	26 (3.94%)
Elaboration	9 (6.82%)	38 (28.79%)	3 (2.27%)	20 (15.15%)	20 (15.15%)	90 (13.64%)
Opinion	22 (16.67%)	19 (14.39%)	74 (56.06%)	3 (2.27%)	2 (1.52%)	120 (18.18%)
Justification	45 (34.09%)	22 (16.67%)	18 (13.64%)	10 (7.58%)	27 (20.45%)	122 (18.48%)
Summary	1 (0.76%)	11 (8.33%)	1 (0.76%)	45 (34.09%)	4 (3.03%)	62 (9.39%)
Main idea	0	2 (1.52%)	0	24 (18.18%)	0	26 (3.94%)
Reword	0	1 (0.76%)	1 (0.76%)	17 (12.88%)	0	19 (2.88%)
Demonstrating knowledge	3 (2.27%)	13 (9.85%)	12 (9.09%)	16 (12.12%)	7 (5.30%)	51 (7.73%)
Choosing a topic	20 (15.15%)	18 (13.64%)	12 (9.09%)	17 (12.88%)	16 (12.12%)	83 (12.58%)
Taking a stance	53 (40.15%)	3 (2.27%)	24 (18.18%)	0	7 (5.30%)	87 (13.18%)
Argumentation	33 (25.00%)	4 (3.03%)	3 (2.27%)	0	5 (3.79%)	45 (6.82%)
Content	5 (3.79%)	21 (15.91%)	4 (3.03%)	13 (9.85%)	8 (6.06%)	51 (7.73%)
Informing audience	2 (1.52%)	6 (4.55%)	0	3 (2.27%)	23 (17.42%)	34 (5.15%)
Information transformation	6 (4.55%)	11 (8.33%)	8 (6.06%)	8 (6.06%)	24 (18.18%)	57 (8.64%)
Personal experience	0	4 (3.03%)	5 (3.79%)	0	0	9 (1.36%)
Both sides	36 (27.27%)	6 (4.55%)	2 (1.52%)	3 (2.27%)	3 (2.27%)	50 (7.58%)
Subjective	1 (0.76%)	5 (3.79%)	25 (18.94%)	0	0	31 (4.70%)
Objective	3 (2.27%)	8 (6.06%)	0	2 (1.52%)	12 (9.09%)	25 (3.79%)
No addition	0	0	0	11 (8.33%)	0	11 (1.67%)
Other	27 (20.45%)	28 (21.21%)	26 (19.70%)	13 (9.85%)	12 (9.09%)	106 (16.06%)

$p < .001$]. In particular, 27.27% ($n = 36$) of students reported conducting research as necessary for completing the research task, as compared to 3.79% ($n = 5$) of students citing this task element in defining the opinion and summary tasks, and 6.06% ($n = 8$) of students defining the argument task.

3.2.4. Source use

The prevalence of students considering source use as necessary for task completion differed significantly across the five different task assignments, Cochran's $Q(4) = 60.94$, $p < .001$. Post-hoc analyses using McNemar's Test determined that students associated source use more frequently with the research task assignment, than they did with the argument [$\chi^2(132) = 25.41$, $p < .001$], essay [$\chi^2(132) = 11.86$, $p < .001$], opinion [$\chi^2(132) = 35.56$, $p < 0.001$] and summary [$\chi^2(132) = 14.79$, $p < .001$] task assignments. In contrast, students associated source use as a task element significantly less frequently with writing an opinion than they did with writing an essay [$\chi^2(132) = 9.48$, $p < .005$] or a summary ($p < .005$). Specifically, 38.64% of students ($n = 51$) reported source use as a necessary task element when defining a research task, as compared to 5.30% ($n = 7$) to 18.18% ($n = 24$) of students reporting source use as necessary when defining argument, essay, opinion, and summary tasks.

3.2.5. Elaboration

Elaboration as a task element was reported disproportionately frequently across the five task assignments, Cochran's $Q(4) = 49.58$, $p < .001$. Post-hoc analyses using McNemar's Test indicated that elaboration was more frequently reported when students were asked to define the essay task than when they defined argument [$\chi^2(132) = 21.19$, $p < .001$] and opinion tasks [$\chi^2(132) = 31.24$, $p < .001$]. Moreover, students more commonly associated elaboration with summary assignment than they did the opinion, $p < .001$, task assignment. Additionally, students more frequently reported elaboration as a necessary task element for writing a research report than for stating an opinion ($p < .001$). Specifically, 28.79% of students ($n = 38$) reported elaboration as a necessary task element in the case of the essay task, as compared to between 2.27% ($n = 3$) and 15.15% ($n = 20$) of students citing this element when defining opinion, argument, summary, and research tasks.

3.2.6. Opinion

The frequency of reporting opinion as a task element was also significantly different across the five task assignments, Cochran's Q $(4) = 175.33$, $p < .001$. Post-hoc analyses using McNemar's Test determined that students more frequently associated opinion as a task element with the opinion assignment than they did with the argument [$\chi^2(132) = 43.35$, $p < .001$], essay [$\chi^2(132) = 41.07$, $p < .001$], summary [$\chi^2(132) = 65.33$, $p < .001$], and research [$\chi^2(132) = 66.33$, $p < .001$] assignments. Further, students more frequently reported forming an opinion to be a necessary task element in association with the argument task than they did with summary, $p < .001$, or research, $p < .001$, tasks. Additionally, opinion as a task element was cited more frequently in defining the essay assignment than when describing summary, $p < .001$, and research, $p < .001$, tasks. Indeed, 56.06% of students ($n = 74$) reported forming an opinion to be a necessary task element in completing an opinion task, as compared to between 1.52% ($n = 2$) and 16.67% ($n = 22$) of students citing this task element when defining the other four task assignments.

3.2.7. Justification

Cochran's Q determined that the proportion of students defining task assignments as requiring justification varied significantly across the five different task assignments, Cochran's $Q(4) = 37.24$, $p < .001$. Post-hoc analyses using McNemar's Test determined that students associated justification as a task element significantly more frequently with the argument task, than they did with the essay [$\chi^2(132) = 9.88$,

$p < .005$], opinion [$\chi^2(132) = 15.72$, $p < .001$], and summary [$\chi^2(132) = 28.20$, $p < .001$] tasks. Further, students considered the provision of justification to be a necessary task element less frequently when defining summary tasks, than when defining the research [$\chi^2(132) = 8.26$, $p < .005$] task. Specifically, 34.09% ($n = 45$) of students reported justification as a necessary task element for writing an argument, as compared to between 7.58% ($n = 10$) and 20.45% ($n = 27$) of students considering justification to be necessary for completing essay, opinion, summary or research tasks.

3.2.8. Choosing a topic

Across the five task assignments, the percentage of students considering *choosing a topic* to be a necessary task element ranged from 9.09% ($n = 12$) to 15.15% ($n = 20$); however, these percentages were not significantly different across task assignments ($p = .52$).

3.2.9. Taking a stance

Cochran's Q determined that the prevalence of students considering taking a stance to be a necessary task element varied significantly across the five different task assignments, Cochran's $Q(4) = 126.92$, $p < .001$. Post-hoc analyses using McNemar's Test determined that students associated taking a stance as a task element significantly more frequently with the argument task than they did with essay [$\chi^2(132) = 46.17$, $p < .001$], opinion [$\chi^2(132) = 13.75$, $p < .001$], summary [$\chi^2(132) = 51.02$, $p < .001$] and research [$\chi^2(132) = 40.50$, $p < .001$] tasks. Moreover, taking a stance as a task element was more frequently reported in students' conceptions of the opinion task, than in their conceptions of essay ($p < .001$), summary ($p < .001$), and research [$\chi^2(132) = 9.48$, $p < .01$] tasks. Specifically, 40.15% of students ($n = 53$) reported taking a stance as a necessary task element in the case of argument task, as compared to between 0.00% ($n = 0$) to 18.18% ($n = 24$) of students reporting this task element as necessary for composing essay, opinion, summary, or research tasks.

3.3. Research question 3: Task conceptions and performance

For the third research question we were specifically interested in the extent to which defining the argument task in particular ways was associated with performance. Step-wise multiple regression analysis was used. Prior knowledge, attitude certainty, and text condition were controlled for in Step 1 and task elements expected to be associated with improved argument task performance (i.e., taking a stance, providing justification, engaging in argumentation and source use, and considering both sides of issues) were entered as predictors in Step 2. All task elements were entered in a binary fashion. Two models were run predicting students' performance when writing an argument about each of the two topics examined (i.e., flipped classrooms and academic tracking). Additionally, two binary logistic regressions were run to examine the extent to which reporting citation as a necessary task element was associated with actual citation use in students' written responses; and the extent to which considering research and source use to be necessary task components of argument construction was associated with the inclusion of text-based evidence in students' responses. For these binary logistic regressions, prior knowledge, attitude certainty, and text condition were again controlled for in Step 1. Correlations between each task element and students' argument task performance are included in Appendix B.

3.3.1. Flipped classrooms

A regression model was run to predict students' performance on the argument task based on their conceptions of what writing an argument entailed (Step 2), and controlling for prior knowledge, attitude certainty, and text condition (Step 1). However, the overall model predicting argument task performance for the flipped classroom topic was not significant, $p = .24$.

Table 11
Model summary of citation use for flipped classroom topic.

	Predictor Variable	*B*	*SE (B)*	β	*p*
Step 1	Prior knowledge	0.28	0.55	1.33	.60
	Attitude certainty	−0.13	0.14	0.88	.34
	Text condition	−0.64	0.48	0.53	.18
Step 2	Citation	23.26	28393.98	12590401562.5	1.00

Final model: $\chi^2(4) = 9.24$, $p = .06$, Cox and Snell $R^2 = .07$, Nagelkerke's $R^2 = .11$.
Note: Model summary based on final step of the regression.

3.3.1.1. Citation use. A binary logistic regression was run predicting students' use of citation when completing the flipped classrooms argument writing assignment. Prior knowledge, attitude certainty, and text condition were entered at Step 1 and students reporting citation as a necessary task element was entered at Step 2. The overall model was marginally significant, $\chi^2(4) = 9.24$, $p = .06$, with pseudo R^2 measures varying from .07, according to Cox & Snell's R^2, to .11, according to Nagelkerke's R^2. The final model correctly predicted 81.25% of cases. The model summary is shown in Table 11.

3.3.1.2. Text-based evidence. A binary logistic regression was also run predicting students' inclusion of text-based evidence in their responses based on prior knowledge, attitude certainty, and text condition (Step 1) and reporting research and source use as necessary task elements for argument construction (Step 2). The model was overall significant, $\chi^2(5) = 12.17$, $p < .05$. Pseudo R^2 measures ranged from .09 (Cox and Snell) to .14 (Nagelkerke's R^2). The model accurately classified 78.13% of cases as including text-based evidence in their responses, or not. Although the model was significant overall, none of the individual predictors included (i.e., prior knowledge, citing research or source use as a necessary task element) were individually significant. A model summary is included in Table 12.

3.3.2. Academic tracking

A second regression model was run predicting students' performance on the argument task, addressing the topic of academic tracking. Prior knowledge, attitude certainty, and text condition were entered in Step 1 and task elements associated with quality argument formation were entered in Step 2 (i.e., taking a stance, providing justification, engaging in argumentation and source use, and considering both sides of issues). However, the model was not significant, $p = .07$.

3.3.2.1. Citation use. A logistic regression was run predicting students' citation use in composing arguments on the topic of academic tracking. Prior knowledge, attitude certainty, and text condition were entered in Step 1 and students' reporting citation use as a task element necessary for argument composition was included in Step 2. While the overall model was not significant ($p = .12$), Step 2 in the model was significant, $\chi^2(1) = 7.05$, $p < .01$, reflecting the contribution of students' conceptions of argument tasks as requiring the use of citations to actual citation use. A model summary is presented in Table 13.

Table 12
Model summary of text-based evidence for flipped classroom topic.

	Predictor Variable	*B*	*SE (B)*	β	*p*
Step 1	Prior knowledge	0.16	0.58	1.18	.78
	Attitude certainty	0.02	0.13	1.02	.88
	Text condition	0.46	0.46	1.58	.32
Step 2	Research	19.77	13390.47	384613015.97	1.00
	Source use	19.82	10372.97	405403224.73	1.00

Final model: $\chi^2(5) = 12.17$, $p < .05$, Cox and Snell $R^2 = .09$, Nagelkerke's $R^2 = .14$.
Note: Model summary based on final step of the regression.

Table 13
Model summary of citation use for academic tracking topic.

	Predictor Variable	*B*	*SE (B)*	Exp(β)	*p*
Step 1	Prior knowledge	−0.10	0.50	0.91	.84
	Attitude certainty	−0.12	0.14	0.89	.39
	Text condition	0.05	0.47	1.05	.92
Step 2	Citation	22.98	28409.84	9507462452.10	1.00

Final model: $\chi^2(4) = 7.37$, $p = .12$, Cox and Snell $R^2 = .06$, Nagelkerke's $R^2 = .09$.
Final step: $\chi^2(1) = 7.05$, $p < .01$.
Note: Model summary based on final step of the regression.

3.3.2.2. Text-based evidence. A final binary logistic regression was run predicting students' inclusion of text-based evidence in their responses to the academic tracking topic based on prior knowledge. attitude certainty, and text condition (Step 1) and their reporting of research and source use as necessary task elements for argument construction (Step 2). The model was overall significant, $\chi^2(4) = 16.90$, $p < .01$, with pseudo R^2 measures ranging from .12 (Cox and Snell) to .19 (Nagelkerke's R^2). The model correctly classified 78.91% of cases as including text-based evidence in their arguments or not. However, the only individually significant predictor in the model was prior knowledge, β = 3.31, $p < .05$. See Table 14 for a model summary.

4. Discussion

This study examines how students conceptualize five common academic task assignments. Specifically, we asked students to define what they considered writing an argument, summary, essay, opinion, and research report to entail (Research Question 1). Based on these conceptions, we identified important similarities and differences in how students conceptualized each of these tasks (Research Question 2). Moreover, we found that differences in students' conceptions of argument tasks, in particular, meaningfully contributed to task performance (Research Question 3). This study contributes to the literature on multiple text use in at least three ways. First, in response to the theoretical importance placed on students' subjective conceptions of task assignments in theoretical models of multiple text use (Rouet & Britt, 2011; Rouet, Britt, & Durik, 2017), we focus on these. In particular, we are among the first studies, to our knowledge, to explicitly ask students how they define a variety of common academic assignments. Prior work has forwarded the claim that students and faculty, in particular, often have mis-aligned conceptions of what various task assignments entail (Leckie, 1996; Valentine, 2001; Zhu, 2004). We are among the first to provide empirical evidence for this contention. Likewise, researchers note that students' poor performance on argument tasks may be explained, at least in part, by their impoverished argumentation schema (Anmarkrud et al., 2014; Reznitskaya et al., 2001); in examining how students define argument task, we, indeed, demonstrate that this is the case.

Second, we document the focal aspects that students consider when asked to respond to a variety of academic tasks. In doing so, we further

Table 14
Model summary of text-based evidence for academic tracking topic.

	Predictor Variable	*B*	*SE (B)*	Exp(β)	*p*
Step 1	Prior knowledge	1.20	0.57	3.31	.03
	Attitude certainty	0.26	0.14	1.30	.06
	Text condition	0.58	0.47	1.78	.22
Step 2	Research	19.73	13880.12	369616077.21	1.00
	Source use	0.00	0.85	1.00	1.00

Final model: $\chi^2(5) = 16.90$, $p < .01$, Cox and Snell $R^2 = .12$, Nagelkerke's $R^2 = .19$.
Note: Model summary based on final step of the regression.

elucidate long-standing findings in the field suggesting that some tasks (e.g., argument construction) facilitate the comprehension and integration of multiple texts, more so than other tasks (e.g., essay or summary writing, Bråten & Strømsø, 2009; Wiley & Voss, 1999). In particular, we show evidence that students do indeed view argument tasks as requiring "knowledge transformation," or the taking of a stance and provision of justification, as compared to tasks, like essays and summaries, that are focused on "knowledge telling" (Scardamalia & Bereiter, 1987; Wiley & Voss, 1999). Moreover, our results suggest that asking students to compose research reports may facilitate learners' engagement in research, source use, and citation – activities that are oftentimes expected when students are asked to complete a variety of academic tasks, including to compose arguments, essays, and research reports. In capturing how students conceptualize a variety of task assignments, we offer insight into how writing tasks may be better introduced to learners.

Third, we empirically demonstrate the link between students' subjective task conceptions and performance, as is only theorized in the MD-TRACE. Indeed, students' task conceptions were associated with argument task performance, even after attitudes and prior knowledge were controlled for. We believe this represents a key finding, indicating that enriching students' task conceptions may correspond to better performance on a variety of academic tasks.

4.1. Research question 1. Students' conceptions of common academic task assignments

The first research question asked students to define what they considered common academic task assignments to require. Students seemed to have robust schema for some of these task assignments and not others. While in some cases students cited a variety of task elements in association with each assignment (e.g., research and argument task), in other cases students had limited and recursive understandings of what particular assignments required (e.g., defining opinion tasks only as requiring them to provide an opinion). Also notable was the diversity in how some of these task assignments were conceptualized across students.

Students most frequently defined argument tasks as requiring choosing a side in a controversy (40.15%). A substantial, although more limited number of students (34.09%) also considered arguments to require some degree of evidence or justification. Notably, only 27.27% of the sample considered argumentation to require dialogic reasoning or the consideration of both sides of an argument. Indeed, students' engagement in case-building or one-sided argumentation, to the exclusion of conflicting points of view, has been cited as a limitation of argumentation skills, even among undergraduate students (Bråten, Ferguson, Strømsø, & Anmarkrud, 2014; Iordanou, Kendeou, & Beker, 2016; McCrudden & Barnes, 2016; Wolfe & Britt, 2008). Interestingly, engaging in research, using sources, or providing citations were task elements not commonly cited in association with the argument task. Collectively, these three task elements were cited only by 18.19% of the sample. Although students did report a need to provide justifications for claims, within the context of argument construction, such justifications were discussed only generally, rather than in terms of engaging in specific research or information use. This suggests that in assigning argument tasks to students, greater specification of the type of justification required and the extent to which such justifications ought to be based on research, may be necessary.

Understandably, students conceptualized the essay task as the most ambiguous or general of the five task assignments. As a result, students most frequently reported writing and elaboration as task elements necessary for essay writing. The writing category corresponded to determinations of length and format, while the elaboration category reflected students' general statements about the need for description, discussion, or explanation. As such, both the writing and the elaboration categories may be considered to be generic approaches to defining

task assignments, just as essay task assignment leaves students with a great deal of uncertainty regarding what may be expected of them. Students' limited conceptions of essay tasks suggest that when instructors provide general essay prompts to students, the expected task demands (e.g., in terms of research, source use, or argument formation) need to be explicitly outlined to a greater extent (Norton, 1990). In particular, given the frequent use of essay prompts in classroom and standardized assessment students may need more direct training in what quality essay composition entails (e.g., Bridgeman, 1991; O'Hara & Sternberg, 2001).

Opinion tasks were defined overwhelmingly by students as requiring the expression of personal views or perspectives. This is reflected in 45.45% of students' task conceptions coded into the recursive category and 56.06% of students' task conceptions coded as requiring opinion formation. Conversely, very few students considered opinion tasks to require research (3.79%), source use (5.30%), or any kind of justification (13.64%). Collectively, these conceptions seem to indicate that students consider opinion tasks to require only the expression of personal preference, sans the provision of any research, justification, or support.

Summary tasks were primarily defined recursively, as ones that ask students to summarize information; such descriptions were characteristic of 34.85% of students' conceptions. When not defined recursively, summary tasks were defined by students according to their length, corresponding to the writing category, and as *brief overviews*, coded into the summary category. As such, summaries were conceptualized rather distinctly or idiosyncratically by students, as compared to the other task assignments, perhaps corresponding to the prominence of summary tasks across the curriculum (Hidi & Anderson, 1986; Chang, Sung, & Chen, 2002; Leopold & Leutner, 2012). At the same time, the features that students reported in defining summary tasks (e.g., length) were rather superficial in nature. Given the frequency with which summary tasks are used in academic settings to assess comprehension, we hoped that students would describe summaries in more nuanced ways and in ways better aligned to the literature. Duke and Pearson (2009) define summaries as requiring the identification and synthesis of important information, presented in a larger text, to create a new, stand-in text; such a conception of summary tasks is quite far from students' conceptions of these assignments.

Research reports were commonly defined recursively, as requiring research. Additionally, research reports were defined in terms of writing, or length and format, and source use. Moreover, conducting research or gathering information was reported as a necessary task element by 27.27% of students, more so than for any other task. Although not cited by the majority, 17.42% of students considered research reports to also require citation, making it the task most associated with citation use.

4.2. Research question 2. Task elements across task assignments

Research Question 2 examined the prevalence of various task elements across students' conceptions of the five different task assignments (i.e., argument, essay, opinion, summary, and research reports). We begin by discussing those task elements that were commonly reported across task assignments. Then, we discuss task elements that emerged somewhat uniquely in association with particular task assignments, and not others. Finally, we discuss those task elements that we expected to emerge in students' task conceptions but that, nevertheless, were not strongly manifest.

4.2.1. Task elements common across task assignments

A number of task elements were found to be commonly reported across task assignments. For one, writing was an element frequently cited across tasks. This demonstrates that students viewed composition format and length as important components of assignment completion. Nevertheless, concerns about writing represent fairly superficial

understandings of task, demonstrating little concern for quality of content. The need for justification was also a task element frequently reported across the five different task assignments. This was an encouraging finding, indicating that students felt the need to provide evidence in support of claims or considered this to be expected, when completing common academic task assignments. Nevertheless, this justification category included only students' mentions of the need to provide evidence, in a very general sense. This suggests that students may need better developed schema for what kinds of evidence may be appropriate across domains and in response to different types of tasks. The need for students to more specifically understand appropriate evidence types is further indicated by the relatively limited extent to which the need for research, source use, or citation were reported across students' task conceptions.

4.2.2. Task elements particular to certain task assignments

Several task elements coded for were associated with conducting research. These included task elements associated with research, source use, and citation. All three of these categories were disproportionately reported in association with the research task. While this may be expected, it is interesting to note that research and source use were not particularly associated with argumentation, despite this task assignment previously having been found to result in improved and more integrative source use, at least within the context of multiple text tasks (Wiley & Voss, 1999). This suggests that argument tasks may not spur information use, in and of themselves, relative to other task assignment, like writing a research report. Nevertheless, argument tasks may facilitate improved information use, when a set of texts are provided, through students' conceptions of argument assignments as requiring justification.

We were also particularly interested in the role of argumentation in assignment composition. Engagement in argumentation corresponded to task elements like taking a stance, argumentation, and considering both sides of issues. As may be expected, these task elements were predominantly cited in association with the argument task assignment. Comparing these results to task elements cited when participants were asked to define what writing a research report entailed, these task elements were largely absent. Examining argument tasks vis-a-vis research reports suggests that students understand research reports to require information use in a neutral or objective fashion, whereas argument tasks ask that students choose and defend a personal position. This provides evidence for students' conceptions of tasks as either personal or more objective in nature, as suggested by Wiley and Voss (1999). Of course, in the ideal, task instructions might be ones that both encourage students to engage in information use and to develop personal perspectives or stances on information.

We also compared students' conceptions of argument tasks vis-à-vis summary tasks, as these have been the task assignments most commonly compared in prior research (Gil et al., 2010a; 2010b; Wiley & Voss, 1999). Students' conceptions of each of these tasks make clear why differences in performance emerge in response to each of these two task assignments. In particular, while summary tasks were fairly narrowly defined as requiring the provision of a general overview or the identification of important ideas, argument tasks were more comprehensively defined as requiring students to adopt a particular stance on a topic, provide justification for that stance, and, even, potentially, consider conflicting information. In this way, argument tasks seem to be conceptualized by students as requiring more active or deeper-level processing, than summary tasks, which seemed to require only the reproduction and reduction of information.

4.2.3. Task elements infrequently reported

In addition to students' conceptions of particular task assignments, there were a number of task elements that, to our surprise, were cited relatively infrequently, relative to their importance in composing quality responses to a variety of task assignments. To start, students

rarely considered citation use to be necessary for task completion. Even when defining the research task assignment, the need to include citations was identified by only 17.42% of respondents. On a related note, in this study we were specifically interested in determining which task conditions spurred students to engage in multiple text use, including in the evaluation and integration of multiple texts. However, students rarely considered any of the five task assignments they were asked to define as requiring the use of multiple texts or the critical evaluation or analysis of information. Although one student did define composing a research report to require: *gather[ing] a lot of facts from multiple credible sources*, such task conceptions were rare. Only five students reported the need to include reliable sources to be a part of completing research tasks; while no students did so when defining argument, essay, and opinion tasks.

Another deficit in students' task assignment conceptions was considering tasks to require deep-level engagement with information, including the analysis, synthesis, or critical evaluation of information. Students' reports of task assignments requiring any degree of deep-level processing were coded into the *information transformation* category. This included responses such as: *this would cause me to think about the topic a lot more and come up with more in-depth questions.* However, responses associated with information transformation only constituted 8.64% of the conceptions reported, across task assignments. This seems to indicate students' desires to engage with information not only to a limited extent, but also at a superficial level.

4.3. Research question 3. Task conceptions and performance

Research Question 3 examined students' conceptions of argument tasks and the association between students' conceptions of argument tasks and performance. The argument task, specifically, was the focus of investigation because such tasks have been found to confer benefits for students' multiple text integration (Gil et al., 2010a; 2010b; Wiley & Voss, 1999). Across topics, linear regression models predicting overall argument task performance, based on particular task elements cited in students' conceptions of argument tasks, were not significant. Nevertheless, the inclusion of some task elements in defining what writing an argument entailed (i.e., citation use, research and source use) was associated with particular aspects of task performance (i.e., inclusion of text-based evidence; citation use). This discrepancy in findings across specific and more global measures of performance may indicate that the rubric used to holistically score students' argumentative responses was too general to specifically map onto the particular task elements that students identified in their argument task conceptions.

Across topics, binary logistic regression indicated that considering citation use to be a requirement for completing argument tasks was a significant predictor of whether or not students included citations in their responses. This indicates the link between students' task conceptions and argument task performance. In other words, students who a-priori considered argument tasks to require citation use, would then have attended to and considered document information (e.g., author, publisher) during text processing, and would have included such information (i.e., citations) in their written responses. This is an encouraging finding given the number of interventions developed to foster students' sourcing, including engagement in citation use (Brante & Strømsø, 2017; Britt & Aglinskas, 2002; Paul, Macedo-Rouet, Rouet, & Stadtler, 2017; Stadtler, Paul, Globoschütz, & Bromme, 2015). As such, we contribute to the literature demonstrating that prompting students to know that citation is necessary for task completion is a viable strategy for increasing citation use (Paul et al., 2017; Stadtler, Paul, Globoschütz, & Bromme, 2015).

Likewise, conceptualizing argument task as requiring research or source use was associated with students' inclusion of text-based evidence in their written responses. This finding is notable because it demonstrates that particular task conceptions can cue strategy use and behavior during multiple text task completion. In particular,

considering research or source use to be necessary for response composition may have prompted students to more closely attend to text-based information and to include such information as evidence in response composition. Put another way, considering research and source use to be necessary for argument composition may have helped students to set an epistemic standard for the type of evidence that was needed to justify their argument on a controversial topic (i.e., text-based evidence; Greene, Azevedo, Torney-Purta, 2008; Mason & Boldrin, 2008).

This study contributes to the literature in three primary ways. First, from a theoretical standpoint, despite the importance of task assignment or goals in students' performance on a variety of tasks (McCrudden & Schraw, 2007; Rouet & Britt, 2011), no other work, to our knowledge, has examined students' own conceptions of task assignments. In this study, we were able to document important commonalities and differences, as well as limitations, in how students understood common academic tasks. In particular, we contribute further to investigations of the MD-TRACE, with a particular focus on how students' task models may be formed (Rouet, 2006; Rouet & Britt, 2011). We demonstrate here that task models are formed, in part, according to students' task conceptions. Moreover, we are able to tease apart students' conceptions of argument tasks, found to be particularly facilitative for multiple text task performance, vis-à-vis conceptions of summary tasks. Second, results from this study indicate that students have distinct schema associated with different types of task assignments. Notable among these, are that source use, research, and citation use are task elements more associated with research report composition than with any other task assignment. Third, we were able to offer initial evidence that how students conceptualized task assignments is associated with task performance, at least in the case of citation and evidence use. This provides evidence toward a previously only theorized causal mechanism (Rouet & Britt, 2011), whereby students' conceptions of different tasks lead to differences in text processing and task performance.

4.4. Limitations and future directions

Despite the strengths of this study, a number of limitations must be acknowledged. First, the sample targeted in our study were enrolled in educational psychology courses and were pursuing education or education-related majors. Major or domain of study may have affected students' conceptions of the various task assignments examined in this study. The extent to which students majoring in other domains would similarly conceptualize these various academic assignments remains a question and a fruitful avenue for further investigation, particularly given the role of prior knowledge in predicting task performance in the case of the academic tracking topic. Additionally, a variety of writing assignments common in other domains (e.g., laboratory report, persuasive essays, narrative compositions) were not examined in the present study.

In this study, we only examined the association between students' conceptions of argument tasks and performance on an argument construction assignment. Further work is needed to examine the extent to which students' conceptions of the other task assignments examined in this study (e.g., summary, research report) were associated with performance. Moreover, while we demonstrate the relation between task conceptions and performance, more is needed to understand how students' conceptions of various task assignments also dictate differences in text processing and strategy use. There is also a need to generally replicate the findings in this study to understand the stability of students' conceptions of tasks.

Finally, the task formulations that students were asked to define in this study were brief and somewhat decontextualized. The way students conceptualize task assignments presented in more naturalistic settings and with additional parameters specified by instructors is an interesting question for future investigations. Recently, Rouet et al. (2017) proposed the Reading as Problem Solving (RESOLV) model as a framework for understanding how learners conceptualize the contextual features of particular learning situations, or form *context models*. Students' conceptions of contexts are based on five components: the request made, the requester, the audience, the supports and potential obstacles associated with request completion, and the self, or students' conceptions of themselves as learners. While each of these contextual components may be differentially prominent in different learning situations, each of these needs to be systematically investigated, in association with variations in task goals, in subsequent research. As a final point, a number of students in this study defined task assignments recursively, using the same terms as were used in the task assignment (e.g., defining opinion task as requiring the provision of an opinion). Future iterations of this study should modify instructions to explicitly ask students to define task assignments in their own words.

5. Conclusions and Implications

Findings from this study indicate that students have well-developed conceptions of many, but not all, common academic task assignments and that these conceptions are differentiated in association with different tasks. Moreover, students' conceptions of argument tasks, in particular, were associated with performance. These findings suggest that improving students' performance on common academic tasks may be, in part, a question of clarifying expectations for task completion. For instance, it seems to be the case that assignments asking students to compose arguments, may benefit from further instructions asking students to engage in source use. At the same time, evidence shows that students, even at the undergraduate level, have somewhat superficial conceptions of task assignments, for instance focusing on writing length or format. This was particularly the case for commonly used academic assignments like essay writing and summary composition. This suggests the need to enrich students' schema of what common academic task assignments require and, in particular, to habituate students to regularly conducting research as a part of academic writing. It is our hope that findings from this study offer a baseline model that teachers and researchers can use to understand and develop students' task conceptions.

Appendix A

Dr. Andrews
Professor of Education Policy
Duke University
OR
Mr. Matthews
Social Studies Teacher
Question: Should high schools implement academic tracking?
Answer: Based on my decades of research on academic tracking, **I consider tracking to be a highly effective educational practice**.
This conclusion is based on numerous studies analyzing students' academic achievement in high-schools with different tracking policies and dozens of interviews with students and administrators.

1) By the time students reach high school, they typically have well-developed academic goals and interests. While some students may want to take advanced courses, like calculus, in preparation for college, for other students, not intending to pursue math-related fields, such courses may be unnecessary. Tracking ensures a strong fit between students' interests and abilities and academic content.

2) Tracking eases the burden on teachers. Instruction is more difficult when teachers have to cater to a wide range of interests and ability levels. Homogenous classes make planning and instruction much easier. Tracking allows teachers to focus on creating interesting and engaging lessons, rather than on catching up students who are behind or challenging students who are bored.

3) Tracking ensures that all students are academically prepared for the classes they're placed in. Tracking prevents students from needlessly struggling or earning low grades in classes that are too hard for them or from being bored in classes that are too easy for them.

4) Tracking offers particular benefits for high-achieving students. These students, despite their abilities, are often bored by content or frustrated when placed in classes with students who are not as motivated as they are. Tracking allows advances classes to move faster and cover high-level topics, while allowing lower-level classes to move at a slower and more deliberate pace.

Appendix B

	Argument Task Performance	
	Flipped Classroom	Academic Tracking
Recursive	−.03	−.18*
Writing	.06	.08
Research	.02	.25**
Source Use	.04	.05
Citation	−.02	.11
Elaboration	.12	.07
Opinion	.02	.01
Justification	.07	.07
Summary	−.01	.08
Demonstrating Knowledge	−.11	−.10
Choosing a Topic	.06	.06
Taking a Stance	.07	.11
Argumentation	−.01	.00
Content	.15	−.05
Informing Audience	.20*	.11
Information Transformation	.19*	.11
Both Sides	.05	.01
Subjective	.14	.14
Objective	−.07	.02

Note: * $p < .05$; ** $p < .01$; *Main idea, reword,* and *personal experience* were removed from correlation analyses because no students cited these task elements as necessary when completing an argument task.

References

Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. Israel, & G. Duff (Eds.). *Handbook of research on reading comprehension* (pp. 69–90). New York, NY: Routledge.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*(5), 888–918.

Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences, 30*, 64–76. https://doi.org/10.1016/j.lindif.2013.01.007.

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science, 41*(5), 873–894. https://doi.org/10.1007/s11251-013-9263-8.

Brante, E. W., & Strømsø, H. I. (2017). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psyhology Review, 1–27*. https://doi.org/10.1007/s10648-017-9421-7.

Bråten, I., & Strømsø, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47*(1), 1–31. https://doi.org/10.1080/01638530902959646.

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2014). Students working with multiple conflicting documents on a scientific issue: Relations between epistemic cognition while reading and sourcing and argumentation in essays. *British Journal of Educational Psychology, 84*(1), 58–85. https://doi.org/10.1111/bjep.12005.

Bridgeman, B. (1991). *Essays and multiple-choice tests as predictors of college freshman GPA.* Princeton, NJ: Educational Test Service.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction, 20*(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2.

Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction, 19*(1), 13–27. https://doi.org/10.1016/j.learninstruc.2007.12.003.

Chang, K. E., Sung, Y. T., & Chen, I. D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education, 71*(1), 5–23. https://doi.org/10.1080/00220970209602054.

Duke, N. K., & Pearson, P. D. (2009). Effective practices for developing reading comprehension. Retrieved from *The Journal of Education, 189*(1/2), 107–122. http://www.jstor.org/stable/42748663.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010a). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology, 35*(3), 157–173. https://doi.org/10.1016/j.cedpsych.2009.11.002.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010b). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology, 31*(1), 30–68. https://doi.org/10.1080/02702710902733600.

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and mythological directions. *Educational Psychologist, 43*(3), 142–160. https://doi.org/10.1080/00461520802178458.

Hagen, Å. M., Braasch, J. L. G., & Bråten, I. (2014). Relationships between spontaneous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. *Journal of Research in Reading, 37*(S1), S141–S157. https://doi.org/10.1111/j.1467-9817.2012.01536.x.

Hidi, S., & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research, 56*(4), 473–493.

Iordanou, K., & Constantinou, C. (2014). Developing pre-service teachers' evidence-based argumentation skills on socio-scientific issues. *Learning and Instruction, 34*, 42–57. https://doi.org/10.1016/j.learninstruc.2014.07.004.

Iordanou, K., Kendeou, P., & Beker, K. (2016). Argumentative reasoning. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.). *Handbook of epistemic cognition* (pp. 39–53). New York, NY: Routledge.

Kaakinen, J. K., & Hyönä, J. (2008). Perspective-driven text comprehension. *Applied Cognitive Psychology, 22*(3), 319–334. https://doi.org/10.1002/acp.1412.

Kobayashi, K. (2009). Comprehension of relations among controversial texts: Effects of external strategy use. *Instructional Science, 37*(4), 311–324. https://doi.org/10.1007/s11251-007-9041-6.

Kuhn, D. (2010). Teaching and learning science as argument. *Science Education, 94*(5), 810–824. https://doi.org/10.1002/sce.20395.

Kuhn, D., & Moore, W. (2015). Argumentation as core curriculum. *Learning: Research and Practice, 1*(1), 66–78. https://doi.org/10.1080/23735082.2015.994254.

Le Bigot, L., & Rouet, J. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research, 39*(4), 445–470. https://doi.org/10.1080/10862960701675317.

Leckie, G. J. (1996). Desperately seeking citations: Uncovering faculty assumptions about the undergraduate research process. *The Journal of Academic Librarianship, 22*(3), 201–208.

Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction, 22*(1), 16–26. https://doi.org/10.1016/j.learninstruc.2011.05.005.

List, A., & Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educational Psychologist, 52*(3), 182–199. https://doi.org/10.1080/00461520.2017.1329014.

List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist, 54*(1), 20–39. https://doi.org/10.1080/00461520.2018.1505514.

List, A., Alexander, P. A., & Stephens, L. A. (2017). Trust but verify: Examining the association between students' sourcing behaviors and ratings of text trustworthiness. *Discourse Processes, 54*(2), 83–104. https://doi.org/10.1080/0163853X.2016.1174654.

List, A., Grossnickle, E. M., & Alexander, P. A. (2016). Profiling students' multiple source use by question type. *Reading Psychology, 37*(5), 753–797. https://doi.org/10.1080/02702711.2015.1111962.

Mason, L., & Boldrin, A. (2008). Epistemic metacognition in the context of information searching on the web. In M. S. Khine (Ed.). *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (pp. 377–404). New York, NY: Springer.

McCrudden, M. T., & Barnes, A. (2016). Differences in student reasoning about belief-relevant arguments: A mixed methods study. *Metacognition and Learning, 11*(3), 275–303. https://doi.org/10.1007/s11409-015-9148-0.

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review, 19*(2), 113–139. https://doi.org/10.1007/s10648-006-9010-7.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology, 35*(4), 229–241. https://doi.org/10.1016/j.cedpsych.2009.12.001.

McCrudden, M., Magliano, J., & Schraw, G. (2011). The effect of diagrams on online reading processes and memory. *Discourse Processes, 48*(2), 69–92. https://doi.org/10.1080/01638531003694561.

Monte-Sano, C., & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research, 44*(3), 273–299.

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly, 46*(3), 273–304.

Norton, L. S. (1990). Essay-writing: What really counts? *High Education, 20,* 411–442.

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education, 76*(1), 59–92. https://doi.org/10.3200/JEXE.76.1.59-92.

O'Hara, L. A., & Sternberg, R. J. (2001). It doesn't hurt to ask: Effects of instructions to be creative, practical, or analytical on essay-writing performance and their interaction with students' thinking styles. *Creativity Research Journal, 13*(2), 197–210.

Paul, J., Macedo-Rouet, M., Rouet, J., & Stadtler, M. (2017). Why attend to source information when reading online? the perspective of ninth grade students from two different countries. *Computers and Education, 113,* 339–354. https://doi.org/10.1016/j.compedu.2017.05.020.

Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology, 69*(4), 309–315.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Qualtrics Labs, Inc [Computer software]. (2009). Provo, Utah, USA. Retrieved from: http://www.qualtrics.com.

Reznitskaya, A., Kuo, L., Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual Differences, 19*(2), 219–224. https://doi.org/10.1016/j.lindif.2008.11.001.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2–3), 155–175. https://doi.org/10.1080/0163853X.2001.9651596.

Rouet, J. (2006). *The skills of document use: From text comprehension to web-based learning.* Mahwah, NJ: L. Erlbaum Associates.

Rouet, J., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.). *Text relevance and learning from text* (pp. 19–52). Charlotte, NC: Information Age.

Rouet, J., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015.

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. *Advances in Applied Psycholinguistics, 2,* 142–175.

Schraw, G., Wade, S. E., & Kardash, C. A. M. (1993). Interactive effects of text-based and task-based importance on learning from text. *Journal of Educational Psychology, 85*(4), 652–661. https://doi.org/10.1037/0022-0663.85.4.65.

Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer-Supported Collaborative Learning, 2*(2), 191–210. https://doi.org/10.1007/s11412-007-9015-3.

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met.a.ware on the web search of laypersons. *Computers in Human Behavior, 24*(3), 716–737. https://doi.org/10.1016/j.chb.2007.01.023.

Stadtler, M., Paul, J., Globoschütz, S., & Bromme, R. (2015). Watch out!—An instruction raising students' epistemic vigilance augments their sourcing activities. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.). *Proceedings of the 37th annual conference of the cognitive science society* (pp. 2278–2283). Austin: Cognitive Science Society.

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*(8), 1081–1087. https://doi.org/10.3758/BF03206376.

Valentine, B. (2001). The legitimate effort in research papers: Student commitment versus faculty expectations. *The Journal of Academic Librarianship, 27*(2), 107–115. https://doi.org/10.1016/S0099-1333(00)00182-8.

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*(2), 301–311. https://doi.org/10.1037/0022-0663.91.2.301.

Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication, 28*(2), 193–219. https://doi.org/10.1177/0741088311399236.

Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking & Reasoning, 14*(1), 1–27. https://doi.org/10.1080/13546780701527674.

Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication, 26*(2), 183–209. https://doi.org/10.1177/0741088309333019.

Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of Second Language Writing, 13*(1), 29–48. https://doi.org/10.1016/j.jslw.2004.04.004.