

Summer 8-2022

# Science Education Research on Student to Student Discourse: A Literature Review with an Applied Linguistics Lens

Abbee A. Mortensen  
*Portland State University*

Follow this and additional works at: <https://pdxscholar.library.pdx.edu/honorsthesis>



Part of the [Applied Linguistics Commons](#), and the [Discourse and Text Linguistics Commons](#)

**Let us know how access to this document benefits you.**

---

## Recommended Citation

Mortensen, Abbee A., "Science Education Research on Student to Student Discourse: A Literature Review with an Applied Linguistics Lens" (2022). *University Honors Theses*. Paper 1274.

This Thesis is brought to you for free and open access. It has been accepted for inclusion in University Honors Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

**Science Education Research on Student to Student Discourse: A Literature Review with an Applied  
Linguistics Lens**

By

Abbee Mortensen

An undergraduate honors thesis submitted in partial fulfillment of the

requirements for the degree of

Bachelor of Arts

in

University Honors

and

Applied Linguistics

Thesis Advisor

Alissa Hartig

Portland State University

2022

### **Abstract**

Science education researchers have increasingly been using discourse analysis to research small group discussion while doing course related activities. This literature review examined studies from that discourse community to see what analysis methods were used and if they connected to applied linguistics theories or methods. Two of the most common discourse analysis frameworks used were Toulmin's Argument Pattern and grounded coding schemes that looked at the design and specific contents of the discourse. Studies covered will also include education scholars using theories of conversation analysis and normative pragmatics to make a more explicit connection to applied linguistics. Orienting and drawing from applied linguistics can benefit science education researchers in their continuous exploration of methods to measure student learning outcomes through their discourse in small group discussions.

## Introduction

As STEM fields have heeded calls to employ more active learning strategies among primary and secondary education (Bierema et al., 2017), analyzing the discussions students have within these classes has become an important tool for assessing the quality of their learning (Moon et al., 2017). Small group discussions are a part of active learning because they allow autonomy for students and a space to deepen their understanding of science topics along with importance in being able to “talk science” (Bennet et al., 2010), which allows students to learn how to communicate in ways that match the expected discourse of the field. These discussions can be a key place for students to develop scientific argumentation skills and think critically about the course concepts (Rudsberg et al, 2016, p. 710). Some scholars in science education have chosen to research what is happening within these small group discussions, such as looking into using inquiry based approaches in the discussion (Moon et al., 2017; Prince et al., 2018). They look to them to examine aspects such as the quality of arguments students are making (Evagorou & Osborne, 2012; Kulatunga et al., 2013a; Walker & Ziedler, 2007) or if students are responding and building off of each other’s ideas (Barth Cohen et al., 2016; Bierema et al., 2017; Young & Talanquer, 2013).

Previous literature reviews have looked into the existing research on student small group conversations. One example is Bennet et al. (2010), which examined 92 studies that concerned student group discussions of ages 11-18 from 1980-2005. Of the 92, they found that only 24 of them met their criteria for studies that had more in depth research. This criteria was if the study either gave detailed descriptions of what occurred or it evaluated different scaffolding strategies that supported the discussions (pp. 76-77). They noted that the common methods used to analyze their data were either an existing model or developing a unique grounded coding scheme. After their complete analysis of all the studies, they found that a “complex and interacting set of factors are involved in enabling students to engage in dialogue” (p. 91). The “complex and interacting set of factors” speaks to why researchers can

benefit from using applied linguistics research methods. They can benefit from these methods because they were created to address and capture the complexities of language, which can lead to findings that better evaluate students' learning. Overall, Bennet et al. identified that more research is needed to see how using "established discourse analysis techniques" (2010, p. 91) may benefit the study of small group discussions.

Another literature review was done over research on students' discourse by Howe and Abedin (2013). Their research supported some of Bennet et al.'s (2010) findings. In regards to studying group conversations, Howe and Abedin found that qualitative approaches are still most common in "model-based approaches" (2013, p. 345) which are similar to the methods of using an existing framework or grounded coding scheme that Bennet et al. (2010) described. Howe and Abedin (2013) also noted that science disciplines in particular were most represented in studying small group conversations. This could be an indication that science education in particular could benefit from using interdisciplinary techniques to analyze discourse.

### **Discourse Analysis**

Applied linguistics is an inherently interdisciplinary field. It concerns the study of language and communication in order to apply to everyday situations and problems, including social and mental elements. This means it can interact with any other discipline to look at discourse used within their area and also draw and expand on other disciplines' theories (Conrad et al., 2021). This literature review looked at what occurs when science education uses discourse analysis and at times other specific methods associated with applied linguistics. The focus on this discipline was because science education and its sub disciplines (e.g., chemistry education, biology education) are likely to be interested in researching the way students interact with each other peer to peer in a classroom. While many applied linguistic theories can be utilized by education researchers, discourse analysis is especially useful in the case of studying small group discussions. Applied linguists are interested in areas where problems in

communication are occurring and classroom discourse can benefit from exploring its particular language related issues. This is an area where applied linguistics and science education discourse communities have overlap in their research goals, so they can benefit from sharing methods and theories.

This paper looked into education research to review the methods they use to study student small group discourse and addressed the following research questions:

1. When peer to peer discussions of students are studied through discourse analysis, what specific coding methods are used?
2. Do they connect to other theories and subdisciplines of applied linguistics?

These capture an aim of the paper to see how discourse analysis is used in other fields. This also was a response to the observation from Bennet et al. (2010) that more in depth research is needed on this subject.

Discourse analysis, a broad sub discipline within applied linguistics, can be a helpful tool to evaluate students' learning using their classroom discourse. To look at discourse is to examine language in use (Gee, 2015), which can be in a spoken, signed, or written form. Discourse analysis looks at the language used in its context or the environment it occurred in (Hartig, 2021), and is typically connected to multiple voices, perspectives, and intertextual elements. Its origins are interdisciplinary in nature, coming from "philosophy, sociology, linguistics, and literary theory, work in this area has expanded to a wide variety of disciplines, becoming a prominent area of emphasis for research in anthropology, communication, psychology, and education" (Cole et al., 2014, p. 63). Its broad scope contrasts from other sub disciplines of applied linguistics such as phonology, which analyzes the smallest units of language, or psycholinguistics, which focuses specifically on cognitive activity with language. When an applied linguist sets out to analyze discourse, they will specify what method of discourse analysis they are using. This could be a register analysis, which studies typical features of a specific type of discourse,

or corpus linguistics, which uses technology to measure how often a word or phrase occurs and notes the words that are frequently appearing around it.

Education studies that used discourse analysis were examined in this review to highlight the various ways discourse analysis can be applied. It also examined the awareness the studies show to their connection with applied linguistics and related disciplines, such as argumentation and communication. This connection could be shown by naming theories, referencing scholars in these fields, or using a model or framework that has been connected to discourse analysis.

### **Methods**

The goal of this literature review was to showcase the variety of techniques that have been used by science education scholars to research peer discussions. This goal led to a two step process for identifying and collecting the articles within the scope of the research question. The first step was through keyword searches in the Portland State library database. The second step was reviewing reference lists from selected articles. This paper was not an exhaustive review of every method and its frequency.

### **Criteria for Inclusion and Exclusion**

One piece of criteria was that studies needed to include analysis on student to student small group conversations. Studies could have other aspects such as interviewing students or whole group discussions, however only the portions concerning the student to student conversations were used for this review. This student to student requirement included students of any level of experience or skill, such as peer leaders.

Another requirement was that studies must be focused on students in a STEM class. This could include any physical science, life science, psychology, math, computer science, or engineering class. Stokoe (2000) had both psychology and literature classes involved, so it was included because one class met the criteria and the author was in science education.

The last requirement was that the data needed to use a transcript. Transcripts were produced by audio or video recording the discussions and then transcribing them. The study had to present data from the transcripts, but could present findings from additional sources (e.g., interviews or surveys). One example of a study that was excluded is Theobald et al. (2017), which examined when students felt the conversation was “dominated” by another peer and their reactions to it. The researchers also wanted to see how the structure of the activity would influence students’ opinions on it. The data was collected based on student surveys that asked how they felt about their experience and if they felt another student overly dominated the conversation. They also compared scores of a pre and post study test to see how their learning was impacted. This study meets the criteria based on its subject matter, but was excluded because it lacked data from transcripts. If they had created transcripts, more insight into how the different activities impacted conversation and turn taking would be possible. Data also could have been collected on what the speech behavior of a student dominating a conversation looks like, which would have been relevant to this literature review.

### **Article Collection Techniques**

The first step was doing keyword searches using boolean operators in the Portland State Library database. All searches were restricted to articles published in 2000 or later. The keywords were combined using the “AND” boolean operator to have a phrase that involved discussion or conversation of students along with a phrase that signaled active learning methods being used in the classroom. These included the phrases “peer discussion”, “small group discussion”, “STEM”, “science”, “small group conversation”, “student discussion,” “ICAP”, “collaborative learning”, “active learning”, “POGIL”, “toulmin argument pattern”, “conversation analysis”, “discourse analysis”. The main keywords for the “NOT” boolean operator were “second language”, “ESL”, “whole class discussion”. Often, active learning methods will involve a small group discussion or conversation between students while completing an

activity, so it was an effective keyword. “POGIL” would be an example of this because it is a type of active learning framework.

The second step was to examine reference lists of articles that were selected in step one. Articles were identified by titles that suggested they would fit and often included alternate wording for some of the keywords listed in step one. From there, these were also read through with special attention paid to their methods sections to see if they matched the criteria. This was done until saturation was reached, which was measured by a lack of new articles or out of scope content. Out of scope was defined as the article containing research questions that went far beyond the criteria, mainly if studies primarily explored concepts outside of small group discussions. The risk of narrowing the scope also occurred when reviewing reference lists. An example of narrowing the scope would be solely identifying articles using one research method, such as Toulmin’s Argument Pattern (TAP). This literature review aimed to highlight the variety of methods used to analyze small group discussions so it did not focus solely on TAP, but did include a few articles to illustrate how it is applied.

### **Assessment of Articles**

When a potential article was identified, it was read through with close attention paid to the methods section, which would contain most of the information to determine if it met the criteria. This process resulted in 18 articles being used in the literature review. Once all the articles were selected, they were sorted into three broad categories based on the analysis method they used. The different methods and coding schemes were assessed to see how they drew from principles of discourse analysis and other applied linguistics theories.

### **Presentation of Article Analysis Methods**

The 18 articles selected for this study were sorted into three categories based on the analysis method used: Toulmin’s argument pattern (TAP), grounded coding methods, and linguistically oriented methods. TAP was used frequently enough to warrant its own category and this review will highlight

studies that used both TAP and another coding scheme in combination. There were 8 studies in this category. Grounded coding had 8 articles. Any article where the researchers created their own analysis framework through a unique coding scheme was categorized as grounded coding, even if they did not explicitly label it as such. Here the review looks to compare and contrast which linguistic elements were being coded for. Two articles were labeled as linguistically oriented and they were examined in depth for their connections to applied linguistics. All the articles are presented in a table format in Appendix A to compare each article's coding and/or analysis method and show direct quotes mentioning discourse analysis or a related analysis method.

### **Toulmin Argument Pattern**

One approach that came up frequently in science education was Toulmin's Argument Pattern (TAP), which was developed by Toulmin (1958) as a way to analyze the structure of an argument. TAP is also referred to as Toulmin's argumentation scheme and Toulmin's model of argumentation, but will be called TAP as Erduran et al. (2004) chose to abbreviate it. This is a form of discourse analysis that focuses solely on discourse where arguments occur (Cole et al., 2014, p. 65). Argumentation can be developed in student conversations as a way for them to learn science content (Erduran et al., 2004, p. 997). TAP's basic model is that all claims must be supported with "warrants" (evidence) and "backing" (reasoning). The TAP model can also consider rebuttals or qualifiers as part of the argument. In the TAP framework, a claim with no evidence or reasoning would be incomplete. Looking at the "completeness" of an argument through TAP is a way to evaluate the quality of students' knowledge through their discourse. In a student small group discussion, the discourse of focus is their conversation and TAP is used to observe if an argument is created by an individual student or by the group collaboratively. TAP can be used as the sole discourse analysis approach or can be combined with other methods.

### ***TAP Alone***

Kulatunga et al. (2013a) used TAP to see how students were co-constructing arguments. TAP allows an argument to be broken down into the pieces of claim, warrant, backing, rebuttal, and qualifiers. When a claim is paired with warrant and backing, this involves the student applying course content to their claim and can demonstrate their understanding of science knowledge. The researchers could then use the transcript to identify if the argument pieces were coming from more than one student. This created a consistent way to identify if students were building off of each other and having complete arguments occur in their conversations. The co-constructed arguments were more likely to be complete than the individual ones. In this way the TAP framework was applied to show that discourse where students can respond to each other's statements can lead to higher quality arguments that rely on engaging with the course content.

One study, Moon et al. (2017) used TAP to see where students were able to evaluate their own arguments and reflect on others' arguments as well. They used discourse analysis to identify this process of "decentering". The TAP framework allowed them to see that decentering often occurs in the rebuttal piece. The researchers looked for evidence in a student's statement that they were interpreting the previous argument. This involved basic discourse analysis as the researchers considered in context the meaning of each rebuttal and if it connected to the meaning of another's claim. For example, if a student provided their own argument as a rebuttal instead of questioning the incorrect evidence, this would not be decentering. The researchers wanted to see if students would verbally engage with the statements of their peers to show recognition of which facts were incorrect in the disputed claim. This demonstrates decentering from a student's own viewpoint and understanding of why something is wrong. This was only possible through analyzing the transcripts to observe where students interacted with each other in conversation.

***TAP Combined With Other Coding Scheme***

Some studies chose to combine TAP with other coding schemes. Prince et al. (2018) chose to use it along with an additional framework rooted in discourse analysis called Inquiry Oriented Discursive Moves (IODM) (Cole et al., 2014, p. 67). It has four codes that analyzed the linguistic design of how the statements were said: revoicing, questioning, telling, and managing. IODM meant the researchers had to categorize the intent behind statements and how their design functioned in the conversation. Price et al. (2018) presented this quantitatively to gather insights from how often each student did either a TAP part of an argument or an IODM code. The researchers also tracked how many conversational turns each student took. The researchers found the speech behavior of “revoicing” as an accessible entry point to the conversation for students. It can use single words to agree with another student’s argument, which may be easier than stating an entire claim. This parallels the co-construction of arguments that Kalutunga et al. (2013a) looked into.

In this same study, Prince et al. (2018) also looked into the power dynamics of the group. They acknowledged that contributing the most parts of an argument or having the most conversational turns did not always lead to holding the most power in the conversation. This is a topic that is of interest to pragmatics too; these findings are supported by Halvorsen and Sarangi (2014), who found that having the most turns does not indicate having the most power in the group (p. 11). Prince et al. (2018) found that those who had the most frequent instances of the speech behavior “telling” held the most social power. They supported this claim by doing a multi-modal analysis of the student’s discourse that was focused on their physical movements. Students with more power received more eye gaze directed at them by their peers and demonstrated actions like grabbing writing utensils out of peers’ hands. The researchers do not explicitly label this multimodal dimension to their analysis, but it is clear that information from the video recordings informed their discourse analysis. This could show that researchers outside linguistics and communication may feel inclined to look into all aspects of language including non-verbal communication.

One study, Kulatunga et al. (2013b), used TAP combined with a grounded coding scheme to see how the types of questions asked by peer leaders influence the structure of the argument. This study claimed that students have deeper learning when they are forming more complete arguments. If certain questions caused students to have complete arguments, then they were positively impacting their learning. TAP allowed them to evaluate the quality of the activity's questions. Using a grounded coding scheme, the statements of peer leaders were sorted into eight categories based on what was said and how it was said. It was found that questions from the peer leader with a direct answer caused students to more frequently give warrants with their claim. Students were more likely to add in the backing piece when the peer leader asked clarifying questions. This study used TAP to see the results of the questions, but also brought in further discourse analysis frameworks as it analyzed the content and design of utterances by coding questions as "open" or having a purpose of "clarifying".

Another study that added a grounded coding scheme to their TAP analysis was Evagorou and Osborne (2012), who looked at how students create arguments with a socio-scientific problem. They sorted arguments into different levels based on the amount of claims and rebuttals each contained, which is based on a modified TAP framework by Erduran et al. (2004). This was a quantitative method to measure the quality of students' arguments. In addition, they commented on what else happened in the students' discourse. They used codes like: exploratory talk, disputing talk, or agreement without debate. They also looked at shorter utterances to track asking questions or reading aloud the data. This discourse analysis framework covered more of the student discourse to further examine the environment the argumentation occurred in.

### ***Critiques of TAP***

Some researchers have chosen to share limitations of TAP they have observed and suggestions to improve its applications. These studies went farther than adding in an additional coding scheme to their study to use in combination with TAP. They chose to comment directly on what aspects were limited or

not covered by TAP. Some chose to even create new analysis frameworks based off of TAP after sharing their reasoning for why TAP alone was not enough.

Jiménez-Aleixandre et al. (2000) is a study often cited by those using TAP. Here the researchers wanted to separate out when students are “doing science” as in making arguments that engage with the course content or “doing school” which is following tasks because they are told to and know it is expected of them in school culture (Jiménez-Aleixandre et al., 2000, p. 767-8). This study does outline how TAP can be applied to analyzing transcripts of student conversations, but also chose to use an additional grounded coding scheme to look at the “epistemic operations” (p. 759) to analyze if students were using their knowledge accurately. The researchers concluded that a variety of tools and approaches are needed to analyze classroom conversation and that TAP is not enough to interpret certain sequences in the discourse. This was their reasoning for bringing in their epistemic framework and could be seen as a recommendation for other science education researchers to create grounded coding schemes or use existing epistemically centered frameworks. This connects to how discourse analysis has such a wide scope that all of its approaches cannot be listed, as new approaches are being created within it to meet the complexities of research needs. Considering this, it is unsurprising that some who chose to use TAP needed additional discourse analysis methods to answer their questions.

Walker and Zeidler (2007) also used TAP for their study of student discussions to see what parts of arguments were in their discourse. They found that large portions of the dialogue could not be coded for as any part of the argument structure. Transcripts are labor intensive to make and provide large quantities of discourse to analyze. Considering this critique, if much of the data doesn't meet the argumentation structure, researchers may be motivated to pursue another coding scheme that can analyze more of the discussion. Walker and Ziedler (2007) also point out that TAP doesn't take into consideration if the argument is factually correct or not. This connects back to Jiménez-Aleixandre et al. (2000) findings that epistemic elements are not being addressed adequately with TAP. Duschl (2007)

brings up that some researchers who respond to this same issue by coding for the appeals of the argument (e.g. appeal to authority) to show what kind of content is happening. Other researchers may react by creating their own coding scheme to formally or informally expand TAP.

One research team in Rudsberg et al. (2016) also recognized the limitations on TAP and chose to create an expanded version of TAP called Transactional Argumentation Analysis (TAA). The researchers wanted to look at both intrapersonal and interpersonal dynamics of student conversations. They used Dewey's theory of transaction (Dewey & Bentley, 1949, as cited in Rudsberg et al., 2016, p. 713) to inform their view that the environment the students are discussing is being created and modified by their conversation and that the learner and the content they are learning is inseparable. Their research question looked at how the group's co-constructed knowledge influences individual students' understanding and arguments about the science content. This made it important to study elements that go beyond argumentation within the group discussion. The theory of transaction also recognizes the need to look at discourse in context, which has been strongly advised by education scholar Robert Duschl (2007). Duschl notes that the entire reasoning episode should be looked at in sequence and that isolating the claims or warrants can make analysis difficult (2007, p. 169). Discourse analysis researchers would strongly agree with this because a primary motivation for examining the entire discourse in context is to look at how the sequences build off each other and interact with the environment the participants are creating.

Rudsberg et al. (2016) had in-depth descriptions concerning how an individual student's arguments and meaning making were influenced. Data was presented on what the student said and which parts of TAP he contributed. The pieces of the argument were then aligned to a code for transactivity. These categories included: extending someone else's reasoning, critiquing their reasoning, completing another student's claim, or counter arguing. All of these look at how the participants' utterances interact with each other. In order to see how students influence each other this added layer

of transactivity to TAP was needed because solely looking for complete arguments did not go deep enough. Each part of the argument must be compared to which specific parts the other students are contributing to get a fuller picture of how students are building knowledge together and changing their own. This is similar to Moon et al. (2017) which looked at the act of decentering through TAP, but this article is able to explain the differences more explicitly through the additional codes. While Moon et al. (2017) labels interactions in a binary framework of decentering or not decentering, Rundgren et al. (2016) instead looks at the entire conversation as a journey of decentering. They examined a similar idea to decentering by seeing how a student is influenced by and in turn influences his peers across the entire discussion. They chose to label this as “empirical analysis” and used discourse analysis to examine how students’ language behaviors build off of and influence each other.

### **Grounded Coding**

While studies that used TAP gave an example of how to use an existing discourse analysis model, some researchers will choose to create their own methods because the existing models may not sufficiently meet their research needs. For example, they may notice a specific phenomenon or pattern that would be overlooked by an existing model. Faced with this, many researchers will develop grounded codes based on common patterns or features they observe from the transcripts. While observing these patterns, codes will be created by the researcher to accurately describe what linguistic elements in their data connect to their research question. The codes in these studies are grounded in the discourse that they are working with (Eckhart, 2009, pp. 72-73) and create a grounded coding scheme that is a form of discourse analysis.

### ***Two Categories of Codes: Design and Content***

The studies that used a grounded coding approach contained codes that can be divided into two overarching categories: design and content. Nielsen (2011, p. 280) states that linguistics looks separately at the “design” of the message (the linguistic structure, how it is being said) and the “specific contents” of a

message (what is being said). Based on this reasoning, the codes in the following studies were able to be categorized as design or content codes. This section will show studies that used content only codes, but there were no studies with only design codes. It will also show studies that contained both design and content codes, and highlight if they separated them from each other or not.

For the purpose of this literature review, codes which label a speech act or way the student is conducting themselves using language or a nonverbal form of communication will be referred to as “design codes”. Examples of codes used were disagreement, asking a question, or clarifying a statement. These acts can be inferred by looking primarily at the structure of the words. When coding for disagreement, looking for words that indicate opposition would be necessary, such as “No” or “I think instead” but the researchers are not coding what exact concept was disagreed on. They are looking at the general behavior of disagreement, regardless of the content and if the students are conceptually correct or not.

Other codes address statements or questions that look at the epistemic content or referents of the utterance. In this literature review these are referred to as “content codes”. The specific content of what the student is saying is central to these codes; the structure used or how they said it is less important. Examples of specific content codes are accurate/inaccurate, procedural talk, or fact-based. All of these codes deal with the meaning or content of what is being said (the message).

It is possible for a content code to include a design code element within it. One code from Evagorou and Osborne (2012), was “questioning data”, which had the design code of asking a question, but it is a content code because it looks into if the message is referencing the data. The researchers can only apply this code by engaging with the content of the question’s message and making sure it involves the data at hand. Design codes cannot engage with the specific content of the utterance this deeply, so if a code involves both aspects this paper views it as a content code.

It is important to consider whether a code is dealing with the design or specific contents because they are different elements of the way people communicate. Applied linguistics chooses to engage with

both of these codes, and would be careful not to conflate the two. For example, when an applied linguistics study examines question types, they will focus primarily on the design of the question. If they choose to comment on the content of the questions in the discourse, then this would be presented separately.

Content codes can be found by using grounded theory to find what concept is being referenced in the message. Without explicitly acknowledging this or being aware of it, there is a risk of not recognizing how personal biases can cause a researcher to misunderstand the intended meaning of the student. Design codes often look at speech acts because they are looking at the linguistic forms being used. Here researchers can be informed by pragmatics to identify them. For example, in Halvorsen and Sarangi (2014), the discourse type of “presenting an option” (p. 10) in a meeting was identified as a conventional speech act. They looked at the design of what gave an utterance the purpose of presenting a choice. Although they gave examples that showed the specific content of these discourse types, it was not the focus of their analysis. In this sense, design codes tend to have a more broad scope. Design codes also have a strength in being more generalizable; for example, most activities would allow for students to be encouraged to do the speech act of “revoicing”. Nielsen (2011) observes that neglecting to analyze the design can leave out important information because the “arguers perform speech acts that are designed to show (rather than tell) that a standpoint has been adequately argued for” (p. 295). With science education’s interest in student argumentation, focusing only on the specific contents of students’ statements can cause researchers to not consider the strategies students are using to further their arguments.

### ***Content Code Only***

Young and Talanquer (2013) was a study that chose to focus only on content type codes. They coded if the talk was related to the science content or other topics such as the procedure for the activity or completely unrelated to the lesson. The course content-related codes were split into two categories: fact-based or meaning making. Meaning making was defined as “constructing, transforming, and applying academic knowledge” (p. 1124). This was of particular interest because it showed a higher level of

engagement with the content compared to repeating facts. Young and Talanquer (2013) also examined how the type of activity influenced the type of statements students responded with. They looked to see which led to the most instances of meaning-making which is similar to the Kulatunga et al. (2013a), who also considered how the activity's format influenced the discourse, but measured this through completeness of arguments. In Young and Talanquer (2013), Activities with a demonstration or analysis component led to the most instances of meaning making among the students. Coding the transcript and comparing it to the activity type allowed for their discourse analysis to show how the nature of the activity itself can influence students' engagement.

Barth-Cohen et al. (2016) was another study that chose to solely use content codes while they looked for examples of co-constructed knowledge. This is similar to looking for meaning-making like Young and Talanquer (2013), but the meaning-making must be done collaboratively among the students. They identified four content codes to cover instances where co-construction of knowledge occurred, which were: adding science content, acknowledgement of ideas, asking a question about science content, and revising ideas. All of these concern the content of the message. For example, acknowledgment of ideas was defined as "statements that contained instances when a student mentions his or her own ideas or a peer's idea" (p. 57). They are not concerned with the design of how acknowledgement is happening, but instead looking at whether the contents refer to a previously stated idea within the context of the conversation. The focus of their study was on what students say when co-constructing knowledge and not on which speech acts lead to it happening, so it is unsurprising their codes entirely concern the specific content.

Repice et al. (2016) looked at classroom discourse in a peer-led team learning setting where the students discuss and work through class activities in small groups. They chose to sort all speech into two categories: regulative and instructional. Regulative statements were about how to do the activity and instructional was talking about course content. With the foundation for the entire coding scheme being grounded in the content of the message, it is unsurprising that all the codes but one were content codes. All

of the regulative codes were content codes looking at refocusing the topic, giving feedback, and meta-communicative and meta-cognitive statements. The instructional codes were: non-explicit procedural, explicit procedural, conceptual explanation, closed question and open question. Closed and open codes for questions did not refer to the structure of the question; they looked at if they narrowed or broadened the scope of the discussion. The explicitness of the procedural codes did have a design element to them, but overall coding for it was grounded in the content of the message. The sole design code was non-elaborate, which meant having few words in an utterance. This may have been added because the short phrases were hard to code for their meaning. With this study primarily focused on the content, it may have been appropriate to not include these statements in their final data.

### ***Separate design and content codes***

One study, Eckert (2009), did choose to separate out design and specific content codes. She chose to use the word behavior to describe design codes such as: shouting out an answer, asking a question, and rephrasing an answer. She also looked further into the behavior of asking a question, which she sorted into content codes that included: procedural, rhetorical, reflective, and verification. To determine if a question is procedural, Eckert (2009) had to consider if the message's meaning involved concepts relating to starting, doing, or ending the activity at hand. Her choice to not put design and specific contents all in one coding scheme could show an awareness that these are different aspects of communication and worth separating when presenting the data.

Bierema et al. (2017) coded the student discourse to look for instances of sense-making, which was defined as "working to make sense of key aspects of the phenomenon" (p.12). Sense-making is similar to the exploration of meaning-making done in Young and Talanquer (2013). The primary difference is that meaning-making more explicitly names actions that count (like "applying") whereas sense-making requires more interpretation and explanation to identify. Both of these seek to see where students are understanding the content, and choosing to use discourse analysis of small group discussions gives them

appropriate data needed to identify them.

Bierema et al. (2017) choose to focus solely on design codes for their first part of the analysis. For them, looking at the speech acts such as narration, agreement, clarification, and disagreement could indicate sense-making. This contrasts Young and Talanquer's (2013) approach to look at the content of students' messages when they engage in meaning-making because Bierema et al. (2017) chose to consider how students chose to design their statements of sense-making. Bierema et al. (2017) also looked into when students were justifying their contributions with some sort of reasoning. They coded these as: interpretations, resource, or grade-focused. All of these were content codes as they looked at what was being said in the speech act of justification. Justification can show what knowledge a student has about science, which may have been why looking into the specific content was useful to the researchers.

Sampson and Clark (2011) choose to use discourse analysis to compare the conversations of low and high performing groups. They addressed multiple aspects of the discourse using different grounded coding schemes. Each one separated out design codes from content codes. First, they coded if the ideas stated were accurate or inaccurate, which was a content code that is solely dependent on if the utterances refer to correct facts or not. Next, they coded using design types on how students responded to ideas, which were: accept, reject, discuss or ignore. Each of these can be considered a speech act and they found that higher performing groups discussed more often. It was worthwhile to explore deeper into the speech act of discussion, so they coded these instances with the specific content codes of: information seeking, exposition, opposition, co-construction. This is an example of a design code being expanded upon using a content code oriented scheme. Lastly, they coded the specific contents, to see if the evidence was rigorous or informal. The look at accuracy and rigor show discourse analysis beyond TAP being needed to examine the "epistemic operations" (Jiménez-Aleixandre et al., 2000, p. 759) of student discourse.

### ***Combined Design and Content Codes***

While the above studies chose to examine design and content codes in separate schemes, others

did not separate them out so explicitly. This meant that they presented one or multiple coding schemes that contained both design and content codes. In this sense, they combined them together to look at the message's content and how it was said.

One study chose to mainly focus on the speech act of dissent. Kelly et al. (2001) tracked this mostly using design codes such as describing the nonverbal behaviors and looking if the dissent contained speech acts like explaining or refuting. In doing so, this study engaged with the multimodal dimensions of the conversation. They presented diagrams of the students' physical positions so readers could understand the physical dynamic between them. They then commented in depth on instances of dissent to describe what had occurred. It was noted how dissent was used to reopen the topic and would often lead to more scientific arguments being constructed. They also looked into the specific content by coding what the "referent" was with a student's utterance. This was a tool to analyze if their knowledge while dissenting was accurate or not. This pairs well as an additional form of analysis to once again look into "epistemic operations" as Jiménez-Aleixandre et al. (2000, p. 759) called for.

Asterhan and Schwarz (2009) analyzed student conversations about natural selection also using multiple coding schemes, with some separation of specific content and design codes. First, they checked through to see how much content from the course was accurately covered in the conversations. Then they analyzed what was described as microlevel dialogue to look at different speech acts that occurred. They describe their method in coding as the "coding of dialog units was based on both the content and the interlocutory function of the contribution, given the dialogical context." (p. 384). Here they acknowledge both content explicitly and also a design element by referring to the function a speech act serves. This combination of design and content codes include: claim, request information, repetition, continuation, support (for an argument), challenge, and elaboration. With their macro level analysis, it mostly contained design codes like: did both the students contribute, symmetry of contributions, was agreement reached, and if pieces were contributed by different students. These (except agreement reached) look at the design

of how each student contributes to the conversation to see if collaboration is truly occurring. The one specific content code was if the key issue of the activity was brought up or not. This study did not choose to separate out design and content codes as much as others did.

### **Linguistically Oriented**

Some articles written by education researchers explicitly oriented themselves with applied linguistic theories. These researchers went beyond a vague recognition that they were analyzing discourse and instead looked to clearly name their techniques. One study, Stokoe (2000), declared that their research method was conversation analysis and conducted their research in line with its techniques. Another study, Nielsen (2011), highlighted elements of linguistics that are important to a deep analysis of discourse and engaged with them. For example, naming different parts of speech and examining how they shape a student's understanding of the content is more of a linguistically oriented exploration than using broad design codes. This is looking at the smaller parts of structure to examine specific words rather than larger scale speech acts. By engaging with such elements, researchers show greater awareness of the interdisciplinary and linguistically oriented nature of researching small group conversations in science courses.

### ***Conversation Analysis***

Stokoe (2000) chose to use conversation analysis techniques to study small group discussions in psychology and literature classes. She explicitly refers to scholars of conversation analysis, like Gail Jefferson and Mary Howe. Stokoe states these methods were chosen because educational discourse occurs within an institution that is heavy on language use and has participants achieve shared goals through conversations (p. 185). Conversation analysis originates in the applied linguistics field and is based on ethnographic approaches that look at ways people construct social norms and power dynamics. For this discipline, the conversation is studied to see how participating in it creates social order (Liddicoat, 2007).

Sacks, Schegloff, and Jefferson referred to “recipient design” which states that people talk in conversations in ways that make sure the other participants understand their meaning- which is what helps make a conversation successful (as cited in Liddicoat, 2007, p.5-6). For conversational analysis on classroom discourse, recipient design describes how students in small groups speak on course content in ways that their peers can understand. If miscommunications seem to be arising, what is happening around those would be worth analyzing. This could be through looking at how speakers take turns speaking or how often they pause in the conversation. Recurring sequences can be found across conversations and often conversation analysis will look in depth at one sequence. For Stokoe’s study (2000), she chose to look at the opening sequences of small group conversations to see how they orient to the “topic of the day” as defined by the teacher. She also looked into what topics the students seem to view as being on task (p. 189).

For opening sequences, Stokoe’s findings showed that most discussions opened up with discussing the logistics of how to do the task, acknowledging a statement, and then a pause. After this sequence, the students began talking about the content related to the task. This matches conversation analysis findings that opening sequences contain “acknowledgement tokens” and pauses (2000, p. 190). The other type of opening sequence involved discussing absent group members, which Stokoe categorized as “talking on the topic” because it is related to the main task and can serve as a transition to it. She stated that this matches with findings from conversation analysis scholar Sacks of a “false first” topic in conversations (2000, p. 192), which is when participants talk about another topic that is related to their intended topic before beginning to talk about the conversation’s true, intended topic.

The elements of the conversation Stokoe discusses do align with what was coded for in studies that used grounded schemes. She discusses “specific content” by explaining the group is referencing how to do the task or talking about absent group members. Design elements are also focused on by noting pauses, lengthening of words, and through categorizing a response as an “acknowledgement”. The

difference is Stokoe is not counting up how often these occur throughout the entire transcript. Instead she qualitatively comments on the two most common opening sequences and gives examples of each. She explains what happened and how the language shaped the discussion while comparing it to conversation analysis studies. The grounded coding approach studies instead chose to present their data with a strong quantitative component to show how often each code happened. As a whole, this resulted in them being less concerned with the sequences that repeatedly occurred in the student group discussions. This is one of the main differences in these two approaches.

Stokoe (2000) also noted instances where students self-identified that they were off-task in their discussions. This included quotes like “anyway we’re drifting...” (p. 196) and “maybe we should get back to the subject” (p. 197), where students observe that they are off task and have an underlying function of getting the group back on task. There was even an instance of a humorous conversation about the subject matter where a student felt the need to say, “it’s all right we’re discussing sport (.) and corporal punishment” but as the laughter continues, Stokoe suggests that this was said because the students felt they were being off-task due to the responses after acknowledging that they were on a tangent (2000, p. 198-99). When other researchers coded for off-task talk, it was determined by them and did not explicitly state any evidence from the students’ conversation to support it. For example, Young and Talanquer (2013) described this process as “Final codes included constructs such as off-task talk (focused on issues not related to the task at hand)” (p. 1125). They did not choose to analyze how the students oriented away and towards the topic at hand. Stokoe justifies exploring this through a conversation analytic approach by saying “Although [students’] treatment of ‘on-’ and ‘off-task’ talk largely concur with educators’ approaches, much of the detail in the production of topical talk would be missed in a more superficial analysis” (2000, p. 200). Simply coding for off-task talk is a more superficial analysis and misses how students get back on task, which is valuable for understanding the efficacy of small group discussions because it can indicate self-motivation and accountability by the students. Overall, Stokoe

calls for more conversation analysis to be used for insights in educational discourse and to understand what topics students are judging to be important in their conversations (2000, p. 200).

### ***Normative Pragmatics***

Nielsen (2011) is in socio-scientific education and used a framework based in normative pragmatics to study student group discussions. Normative pragmatics states that when a participant states an argument, they cause agreement or disagreement from the other participants and the design of the message (how they choose to say it) affects its interpretation and impact (Jacobs, 2000). Nielsen wanted to look at argumentation patterns and his normative pragmatics approach led him to focus on the design of students' arguments and their "specific effects on the recipients" (2011, p. 279). This separated his research intent from the TAP model that solely focuses on documenting the parts of arguments. While TAP focuses on what is occurring in the entire argument, Nielsen instead looked for what patterns appeared in the environment where the arguments were occurring. This environment was created through the students' discourse and could lead to deeper information about what is happening when they form arguments.

Scholars from communication, like Jean Goodwin (2000), along with Frans van Eemerson of argumentation and pragmatics, are mentioned by Nielsen while he orients himself to work within normative pragmatics. Pragmatics is heavily associated with argumentation, all of which overlap and work within the applied linguistics. He chose this framework because it looks at how arguers (who in this case are the students) look to influence others through the design of their statement (2011, p. 279). He looked for linguistic aspects of the message in order to identify patterns that were occurring in the designs and contents of the students' discourse. His research questions involved looking for how arguments combine "science factual and evaluative statements" and how these work and what kind of language occurs with them (Nielsen, 2011, p. 280).

The first step of his analysis was identifying instances where students combined the science and evaluative statements. From this, common strategies students used were identified. To conduct his normative pragmatics analysis he analyzed the transcripts for four different aspects. Speech acts, argumentative indicators, parts of speech, and connection to the theme of the conversation, were all identified (Nielsen, 2011, p. 282). Speech acts are studied within pragmatics. The argumentative indicators were identified through discourse analysis by considering which language prefaced an argument. He also looked in depth at how pronouns, adjectives, and stance adverbs impacted the discourse, which are all parts of speech and require an analysis rooted in grammar to identify. His examination of connecting the students' turns to the overall theme is further discourse analysis to interpret the specific content their message contains. Each aspect he looks at uses an approach grounded in applied linguistics and its subdisciplines.

The findings showed that students would make arguments by using only scientific facts or would combine these facts with a moral appeal statement (Nielsen, 2011, p. 284). The statements that appealed to morals were further analyzed by looking into the parts of speech used in them. He also chose to describe some adjectives as emotive and others as evaluative, categorizing them by part of speech and by purpose. The adjective "enormous" and stance verb "actually" all make the intensity of the argument higher (Nielsen, 2011, p. 285). Here Nielsen looks at grammatical aspects of the discourse to add to his analysis. This differs from researchers who would mention "appeal to emotions" but not explain the specific words that indicate this nor would they consider how different parts of speech serve specific purposes in a conversation.

Throughout the study a few lines of the transcript are discussed for a thorough qualitative presentation of the results, which parallels Stokoe's (2000) presentation of her findings and the structure of many applied linguistic studies on spoken discourse. Nielsen (2011) looks to identify patterns in the structure for socio-scientific arguments, which can be loosely compared to performing a genre analysis.

This looks at the way statements are organized based on their content or “rhetorical organization” to find consistent patterns (Biber & Conrad, 2009, pp. 16-17). Here, Nielsen wanted to look at the specific pieces that go into creating an argument in the socio-scientific genre. To conclude, he emphasized the complexity of socio scientific student discourse which was highlighted by his in-depth approaches. He also reiterated the importance of looking at the design of statements and not only the content that students are speaking on in their discussions (Nielsen, 2011, p. 295), which can be interpreted as a recommendation to engage with linguistic aspects like grammar and speech acts.

### **Conclusion**

This literature review highlighted the analysis and coding methods science education researchers use to analyze student small group discussions. It presented Toulmin’s Argument Pattern (TAP) as a common discourse analysis tool and also various grounded coding methods. Grounded coding schemes were often combined with TAP and the codes fell into two categories: design and content codes. This connects with Nielsen’s (2011) observation that both the way students are discussing science and what content they are referring to is important for research. There were studies that addressed both aspects, which could indicate that other researchers are noticing their importance. Adding grounded coding to TAP demonstrates a desire to engage with both content and design. Many studies who solely used grounded coding also choose to incorporate codes that addressed both the content and design of the statements. This shows some recognition of the importance of both aspects.

While both TAP and grounded coding methods are forms of discourse analysis, some studies chose to draw explicitly from applied linguistics. These were, Stokoe (2000) and Nielsen (2011), which highlighted the ways education scholars can orient their research to applied linguistic theories. Their use of additional applied linguistics tools such as conversation analysis and pragmatics served to deepen the extent of what can be learned. This review showed that discourse analysis is an effective tool for science education researchers and they are continuing to use it to understand the complexities of student small

group discussion. Future review of the literature can look further into orientation to applied linguistics or even collaborations with applied linguistics scholars.

## References

- Asterhan, C. S. C., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer to peer dialog. *Cognitive Science*, 33(3), 374–400. <https://doi.org/10.1111/j.1551-6709.2009.01017.x>
- Barth-Cohen, L. A., Smith, M. K., Capps, D. K., Lewin, J. D., Shemwell, J. T., & Stetzer, M. R. (2016). What are middle school students talking about during clicker questions? Characterizing small-group conversations mediated by classroom response systems. *Journal of Science Education and Technology*, 25(1), 50–61. <https://doi.org/10.1007/s10956-015-9576-2>
- Bennett, J., Hogarth, S., Lubben, F., Campbell, B., & Robinson, A. (2010). Talking science: The research evidence on the use of small group discussions in science teaching. *International Journal of Science Education*, 32(1), 69–95. <https://doi.org/10.1080/09500690802713507>
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.
- Bierema, A. M.-K., Schwarz, C. V., & Stoltzfus, J. R. (2017). Engaging undergraduate biology students in scientific modeling: Analysis of group interactions, sense-making, and justification. *CBE Life Sciences Education*, 16(4), ar68, 1-16. <https://doi.org/10.1187/cbe.17-01-0023>
- Cole, R. S., Becker, N., & Stanford, C. (2014). Discourse analysis as a tool to examine teaching and learning in the classroom. In D. M. Bunce, & R. S. Cole (Eds.), *Tools of chemistry education research* (pp. 61-81). Oxford University Press.
- Conrad, S., Hartig, A. J., & Santelmann, L. (2021) Introduction to the field of applied linguistics. In S. Conrad, A. J. Hartig, & L. Santelmann (Eds.), *The cambridge introduction to applied linguistics* (pp. 3-20). Cambridge University Press.
- Duschl, R. (2007). Quality argumentation and epistemic criteria. In S. Erduran, & M. Aleixandre (Eds.), *Argumentation in science education perspectives from classroom-based research* (1st ed., pp. 159- 173). Springer.

- Eckart, T. M. (2009). *Productive whole-class discussions: A qualitative analysis of peer leader behaviors in general chemistry*. ProQuest Dissertations Publishing.
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPing into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, 88(6), 915–933.  
<https://doi-org.proxy.lib.pdx.edu/10.1002/sce.20012>
- Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching*, 50(2), 209–237.  
<https://doi.org/10.1002/tea.21076>
- Gee, J. (2015). Discourse, small d, big D. *The International Encyclopedia Of Language And Social Interaction*, 1-5. <https://doi.org/10.1002/9781118611463.wbielsi016>
- Goodwin, J. (2000). Comments on 'Rhetoric and dialectic from the standpoint of normative pragmatics.' *Argumentation*, 14(3), 287–292. <https://doi.org/10.1023/A:1007805130030>
- Hartig, A. J. (2021). Approaches to research in applied linguistics. In S. Conrad, A. J. Hartig, & L. Santelmann (Eds.), *The cambridge introduction to applied linguistics* (pp. 21-35). Cambridge University Press.
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325–356.  
<https://doi.org/10.1080/0305764X.2013.786024>
- Jacobs, S. (2000). Rhetoric and dialectic from the standpoint of normative pragmatics. *Argumentation*, 14(3), 261–286. <https://doi.org/10.1023/a:1007853013191>
- Jiménez-Aleixandre, M. P., Bugallo Rodríguez, A., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, 84(6), 757–792.  
[https://doi.org/10.1002/1098-237X\(200011\)84:63.0.CO;2-F](https://doi.org/10.1002/1098-237X(200011)84:63.0.CO;2-F)

- Kelly, G., Crawford, T., & Green, J. (2001). Common task and uncommon knowledge: Dissenting voices in the discursive construction of physics across small laboratory groups. *Linguistics and Education*, 12(2), 135–174. [https://doi.org/10.1016/S0898-5898\(00\)00046-2](https://doi.org/10.1016/S0898-5898(00)00046-2)
- Kulatunga, U., Moog, R. S., & Lewis, J. E. (2013a). Argumentation and participation patterns in general chemistry peer-led sessions. *Journal of Research in Science Teaching*, 50(10), 1207–1231. <https://doi.org/10.1002/tea.21107>
- Kulatunga, U., & Lewis, J. E. (2013b). Exploration of peer leader verbal behaviors as they intervene with small groups in college general chemistry. *Chemistry Education Research and Practice*, 14(4), 576–588. <https://doi.org/10.1039/c3rp00081h>
- Liddicoat, A. (2007). *An introduction to conversation analysis*. Continuum.
- Moon, A., Stanford, C., Cole, R., & Towns, M. (2017). Decentering: A characteristic of effective student–student discourse in inquiry-oriented physical chemistry classrooms. *Journal of Chemical Education*, 94(7), 829–836. <https://doi.org/10.1021/acs.jchemed.6b00856>
- Nielsen, J. A. (2011). Co-opting science: A preliminary study of how students invoke science in value-laden discussions. *International Journal of Science Education*, 34(2), 275–299. <https://doi.org/10.1080/09500693.2011.572305>
- Prince, A. N., Pitts, W. B., & Parkin, D. W. (2018). Exploring power distribution and its influence on the process of argumentation in a POGIL biochemistry classroom. *Journal of College Science Teaching*, 47(4), 92–107. [https://doi.org/10.2505/4/jcst18\\_047\\_04\\_92](https://doi.org/10.2505/4/jcst18_047_04_92)
- Repice, M. D., Sawyer, R. K., Hoglebe, M. C., Brown, P. L., Luesse, S. B., Gealy, D. J., & Frey, R. F. (2016). Talking through the problems: A study of discourse in peer-led small groups. *Chemistry Education Research and Practice*, 17(3), 555–568. <https://doi.org/10.1039/c5rp00154d>

Rudsberg, K., Östman, L., & Aaro Östman, E. (2016). Students' meaning making in classroom discussions:

The importance of peer interaction. *Cultural Studies of Science Education*, 12(3), 709–738.

<https://doi.org/10.1007/s11422-015-9721-5>

Sampson, V., & Clark, D. B. (2011). A comparison of the collaborative scientific argumentation practices

of two high and two low performing groups. *Research in Science Education*, 41(1), 63–97.

<https://doi.org/10.1007/s11165-009-9146-9>

Stokoe, E. H. (2000). Constructing topicality in university students' small-group discussion: A

conversation analytic approach. *Language and Education*, 14(3), 184–203.

<https://doi.org/10.1080/09500780008666789>

Theobald, E. J., Eddy, S. L., Grunspan, D. Z., Wiggins, B. L., & Crowe, A. J. (2017). Student perception of

group dynamics predicts individual performance: Comfort and equity matter. *PloS One*, 12(7),

e0181336–e0181336. <https://doi.org/10.1371/journal.pone.0181336>

Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.

Walker, K. A., & Zeidler, D. L. (2007). Promoting discourse about socioscientific issues through scaffolded

inquiry. *International Journal of Science Education*, 29(11), 1387–1410.

<https://doi.org/10.1080/09500690601068095>

Young, K. & Talanquer, V. (2013). Effect of different types of small-group activities on students'

conversations. *Journal of Chemical Education*, 90(9), 1123–1129.

<https://doi.org/10.1021/ed400049a>

## Appendix A

Table 1 Summarization of Analysis Frameworks

Article	Coding or analysis type	Explicit mention of discourse analysis or other related analysis
Kulatunga et al (2013a)	TAP	“We have presented a modification of the TAP framework for analyzing student discourse involving argumentation in small group settings based on both the strength of the argument and the extent of participation” (p. 1225)
Walker and Ziedler (2007)	TAP	“Through the process of discourse analysis of the student-pair interviews and researcher triangulation, taxonomic categories of fallacious reasoning, conceptions of science, and sample performances of thought were derived.” (p.1390)
Moon et al (2017)	TAP, content codes about decentering	“Analysis of classroom discourse” (p. 831)
Prince et al (2018)	TAP, design codes using Inquiry Oriented Discursive Moves	“Through the discourse analysis approach and the examination of talk...” (p. 102)
Kulatunga et al (2013b)	TAP, design and content codes	“For our setting, an argumentation framework was used to analyze the student discourse in conjunction with the verbal behaviors of the peer leader.” (p. 577)
Evagorou and Osborne (2012)	TAP, design and content codes	None

Jiménez-Aleixandre et al (2000)	TAP, content codes	“using Toulmin’s model of argumentation as a method of analyzing students’ discourses.” (p. 759)
Rudsberg et al (2016)	Transactional Argumentation Analysis	None
Young and Talenquer (2013)	Content codes	Evaluation of the effects of small-group work often relies on the analysis of student talk or discourse at the content and interaction levels.” (p. 1124)
Barth-Cohen et al (2016)	Content codes	None
Repice et al (2016)	Content codes	“to analyze the discourse, we transcribed video recordings.”(p.558)
Eckert (2009)	Separated design and content codes	None
Bierema et al (2017)	Separated design and content codes	None
Sampson and Clark (2011)	Separated design and content codes	“we relied on a two-stage analytical approach based on the verbal analysis methods developed by Chi (1997)” (p. 75)
Kelly et al (2001)	Combined design and content codes	“analyzed the discourse of one group...” (p.147)
Asterhan and Schwarz (2009)	Combined design and content codes	None
Stokoe (2000)	Conversational analysis techniques	The data was transcribed and subsequently analysed using conversation analysis.” (p. 184)
Nielsen (2011)	Design and content codes, normative pragmatics	“three socio- scientific group discussions were subjected to a normative pragmatics analysis.” (p. 281)

