



Lihua Xu^{1,*}, Jan van Driel² and Ryan Healy³



- ² Melbourne Graduate School of Education, The University of Melbourne, Melbourne, VIC 3010, Australia; j.vandriel@unimelb.edu.au
- ³ Clonard College, Geelong, VIC 3218, Australia; Ryan.Healy@clonard.vic.edu.au
- * Correspondence: lihua.xu@deakin.edu.au

Abstract: Classroom communication is increasingly accepted as multimodal, through the orchestrated use of different semiotic modes, resources, and systems. There is growing interest in examining the meaning-making potential of other modes (e.g., gestural, visual, kinesthetic) beyond the semiotic mode of language, in classroom communication and in student reasoning in science. In this paper, we explore the use of a multi-layered analytical framework in an investigation of student reasoning during an open inquiry into the physical phenomenon of dissolving in a primary classroom. The 24 students, who worked in pairs, were video recorded in a facility purposefully designed to capture their verbal and non-verbal interactions during the science session. By employing a multi-layered analytical framework, we were able to identify the interplays between the different semiotic modes and the level of reasoning undertaken by the students as they worked through the tasks. This analytical process uncovered a variety of ways in which the students negotiated ideas and coordinated semiotic resources in their exploration of dissolving. This paper highlights the affordances and challenges of this multi-layered analytical framework for identifying the dynamic inter-relationships between different modes that the students drew on to grapple with the complexity of the physical phenomenon of dissolving.

Keywords: multimodality; discourse analysis; social semiotics

1. Introduction

Research on the semiotic mode of language as a social action and cultural resource in communicating meaning in classrooms has a long tradition [1,2]. Many studies of science classroom discourse, for instance, have focused on spoken and written language (e.g., [3–5]). In recent years, there has been an increasing interest in examining the meaning-making potential of other semiotic modes (e.g., gestural, visual, kinesthetic) in science classroom communication (e.g., [6,7]), and consideration has been paid to the roles of multiple and multimodal representations in students' meaning-making of science concepts [8–10]. Therefore, classroom discourse is increasingly being recognized as multimodal rather than simply linguistic, through the orchestrated use of different semiotic modes, resources, and systems [11].

This expanded concept of classroom discourse presents both theoretical and methodological challenges for researchers who are interested in understanding how teaching and learning takes place through this complex ensemble of multiple and multimodal representations in a classroom settings, such as a science classroom. Research in this area explores the nature of this coordination of multimodal semiotic resources in student reasoning about science phenomena and to understand how the different modes interact in discourse to generate meanings [12]. In this paper, we discuss the theoretical and methodological challenges that emerged from recent attempts to investigate multimodal reasoning in



Citation: Xu, L.; van Driel, J.; Healy, R. A Multi-Layered Framework for Analyzing Primary Students' Multimodal Reasoning in Science. *Educ. Sci.* **2021**, *11*, 758. https:// doi.org/10.3390/educsci11120758

Academic Editors: Kok-Sing Tang, Kristina Danielsson and Angel M. Y. Lin

Received: 30 August 2021 Accepted: 16 November 2021 Published: 24 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). science classrooms. We propose a multi-layered framework for analyzing student multimodal reasoning. The discussion in this paper will be grounded in an empirical study that aimed to investigate the types of science tasks and classroom interactions that facilitated Grade 5/6 students' reasoning during an open inquiry into the physical phenomenon of dissolving in a primary classroom.

Drawing upon a multimodal semiotic approach for the analysis of classroom video data, this paper addresses the following research question: "what are the affordances and challenges of a multi-layered analytical framework for identifying the dynamic interrelationships between different modes involved in student reasoning about dissolving?" Our focus is to interrogate the affordances of this multi-layered framework, for identifying ways in which the students used different modes to reason and generate meanings about the process of dissolving a solid substance into a liquid, a phenomenon that is not directly observable but can only be understood through inferences.

2. Literature Review

2.1. Video and Multimodal Analysis of Classroom Discourse

The advancement in research into naturally occurring social interactions has been facilitated by the latest innovations in video technologies [13]. Video methods have been developed at an unprecedent rate in educational research in recent years [14,15]. Video data has been identified as a 'real-time sequential medium', which provides a 'fine-grained multimodal record' and is 'durable, malleable, and sharable' [16] (p. 4), which allows for a more precise, complete and fine-grained analysis of human learning, behaviors, and practices [17,18]. The increasing affordability and accessibility of video technology has given rise to the emergence of new methodologies, such as Multimodal Discourse Analysis (MDA), an interdisciplinary field of study that brings discourse analysis and multimodality studies together to understand the design, production, distribution, and interactions of multimodal resources in social settings [19–21]. Research in the field of MDA tends to draw upon the social semiotics approach developed by Halliday and others [21], which adopts a functional approach to meaning making by investigating how various semiotic resource systems have evolved to enable us to perform by making particular kinds of meanings [22]. The term "semiotic resources", as defined by O'Halloran [23] and others [24], refers to the modes of meaning making. These entail both material and immaterial conceptual resources, which are 'realized in and through modes' [24] (p. 71). A semiotic mode is 'a socially organized set of semiotic resources for making meaning' [24] (p. 71). Examples of modes include images, writing and speech. For the clarity of the paper, we will use the term "semiotic resources" to refer to material and conceptual resources for meaning making, whereas we use "modes" to describe the material form(s) of a semiotic resource. For example, a drawing from a student can represent a mixture of modes as a semiotic resource, including both as an image and writing.

2.2. Multiple Representations and Multimodalities in Science Education

The application of a social semiotics approach to discourse analysis can be found in an increasing body of research that investigates the use of multiple representations and multimodality in teaching and learning science (e.g., [6,7,25,26]). A number of previous studies have focused on the use of multiple representations and multiple modes by teachers to support student learning [27–29]. Research in this area demonstrates affordances of different semiotic modes. Examples of modes include image, speech, gesture, action, and so on [20]. Ainsworth [30] emphasized the affordances of different semiotic modes in science teaching, and the need for teachers to consider both students' needs and representational characteristics in designing and delivering learning experiences. The affordances of semiotic modes were illustrated by empirical studies that revealed the roles played by different semiotic modes in science teaching. For example, Márquez, Izquierdo and Espinet [31] applied Systemic-Functional Grammar (SFG) in an analysis of the specialized functions of semiotic modes used by a teacher in teaching the water cycle concept. They demonstrated that, while speech introduces and identifies entities, gestures allow students to locate them in a dynamic process. Visual modes, such as diagrams, facilitate the development of functional mechanisms for the construction of water circulation explanations. As Tversky [32] pointed out, diagrams are ideal for conveying structural organizations by drawing upon people's experience of interpreting special relationships. A more recent study by Moro, Mortimer and Tiberghien [7] investigated how two teachers utilized a range of embodied comparison and an angle of embodied

semiotic modes such as speech, gestures, gaze and proxemics to give meaning to scientific knowledge in the classroom. Their analysis demonstrated the ways in which teachers were able to relate the concrete and abstract aspects of a science concept (e.g., light diffraction) through employing the embodied semiotic modes in a creative and coherent manner. Research in this area also recognizes the need to coordinate multiple modes across scientific and everyday knowledge, and across personal interests to facilitate student

scientific and everyday knowledge, and across personal interests to facilitate student meaning making in science [29]. Kress, Jewitt, Ogborn, and Tsatsarelis [6] documented the complex ensemble of semiotic modes (e.g., image, gesture, speech, writing, models, spatial and bodily movements) brought together by science teachers to construct particular scientific meanings. Employing a social semiotics perspective, Kress and others [6] argued that knowledge construction in science involves dynamic and transformative sign-making from one mode of representation to another and each mode played a particular meaning-making function. Similarly, Tang, Tan and Yeo [33] revealed students needed to construct connections between multiple modes, including verbal, visual, gestural and mathematical modes, in order to understand the work-energy concept.

In this paper, we explore the affordances of a multi-layered analytical framework for investigating students' multimodal construction of representations as they attempted to understand and explain their thoughts about the observed physical phenomenon of dissolving. Through a description of the multiple layers of the analytical process, our intention is to illustrate how different units of analysis allow the researchers to gain insights into the different aspects of the focused multimodal phenomenon (i.e., student reasoning about dissolving), and to interrogate methodological decision making and its consequences when analyzing student reasoning in science.

3. Research Design

3.1. Procedure, Sample

The study reported in this paper was conducted in a laboratory classroom. This particular classroom is equipped with 10 wall and ceiling mounted video cameras with zoom and tilt capacity, and eight radio microphones, controlled from a room with visual access. See Figure 1 for a photo of this purposefully designed classroom.



Figure 1. Photo of the laboratory classroom.

A class of grade 5/6 students (24 boys) from a government primary school participated in a session with their teachers in this classroom, focusing on a science topic. The multiple cameras and microphones meant that 10 video tracks of about 50 min were generated for each session, with a camera focused on each of the six tables to separately capture a video and audio recording of two pairs of students. Thus, the data available for this analysis consists of a continuous visual and audio record of 12 pairs of students for an hour, including a teacher-led plenary introduction and a conclusion of the session. Two ceiling mounted cameras were also employed to provide a separate view of students from above to capture the progress of student discussions and gestures, in conjunction with writing, drawing and their manipulation of equipment and models. In addition, high-definition cameras were used to take pictures of artefacts produced by the students (i.e., worksheets, models, whiteboard drawings).

3.2. Tasks

The tasks for the science session were designed by the research team in consultation with the classroom teacher. Prior to the sessions at the laboratory classroom, the research team attended the participating school and joined a planning meeting with the teachers to discuss suitable topics for the session. Open ended tasks were developed which involve student investigations of dissolving (i.e., of icing sugar in vinegar) and of generating their own representations to make sense of their observations. During the session, the students worked in pairs to investigate this phenomenon. A worksheet was provided to guide them their investigation (see Figure 2). The design of the worksheet follows the Prediction-Observation-Explanation model proposed by White and Gunstone [34]. They were also provided with resources, such as white boards, markers, playdough and toothpicks for them to represent their thoughts on what might be happening during the investigation. The session started with a teacher-led introduction of the investigations and concluded with a whole class discussion during which the teacher asked students to share and explain some of their representations.

Physical and Chemical Change

Team	Memi	her's	names:	
1 cum			mannes.	

÷	Investigation 1: Icing Sugar and Vinegar			
	What do you need?	What do you predict will happen?	What do you observe?	Can you explain what happened?
	One balloon	You may write down or draw your predictions.	Write down as many observations as you can	Imagine if you have a magnifying glass that can be
	 One bottle with 50 ml white vinegar 		(including what you can see, smell, or feel).	zoomed in millions of times, what you might be able
	One funnel			to see?
	 One container with 5 grams of icing sugar 			
	 Playdough 			 Make some annotated drawings OR construct
	Toothpicks			models using playdough and toothpicks to show what
	 Whiteboard 			you might be able to see using such a magnifying
	 Whiteboard markers 			glass:
				1) before adding icing sugar
				2) just after adding icing sugar
	What are you going to do?			3) after shaking for 30 seconds
	1. Open one bottle, smell the vinegar (carefully).			4) after a while, when nothing seems to happen
	2. Add the icing sugar into the balloon using the funnel.			anymore.
	3. Place the balloon on top of the bottle (Do not tip the			
	icing sugar into the bottle, see picture below)			Please use the whiteboard and markers for drawings.
	A			
				2. Can you use the drawings/models to explain some
				of your observations in the 3 rd column?
	4. Before tipping the icing sugar from the balloon into			
	the vinegar, make a prediction of what will happen in			
	the 2 nd column (on the right).			
	5. Now tip the icing sugar from the balloon into the			
	bottle.			
	6. Observe what happens to the icing sugar when it is			
	added to the vinegar before shaking the bottle. Write			
	down your observations in the 3rd column.			
	7. Shake the bottle gently for 30 seconds.			
	8. Observe what happens inside the bottle and what			
	happens to the balloon. Write down your observations			
	in the 3 rd column.			
	9. Feel the bottle and let the bottle sit for a while until			
	no further changes happening. Write down your			
	observations in the 3rd column.			
	10. Open the bottle, smell the liquid. Write down your			
	observations in the 3rd column.			

Figure 2. Student worksheet.

Studies on student ideas of dissolving have a long history, going back to the seminal work of Piaget and Inhelder [35] who demonstrated that young children tend to think that sugar 'disappears' when it dissolves in water, and that the mass of the solution is equal to the initial mass of water. Research on children's conceptions of scientific concepts in the 1980s revealed that many older students (age 9–15) also tend to think that the mass of a sugar solution is less than the combined mass of water and sugar, for instance, "the sugar will decompose and form a liquid with the water and so will weigh less" [36] (pp. 154–155). Prieto et al. [37] asked students (age 11–14) to show, by means of a drawing, how they imagined a substance that is completely dissolved in water and reported that 44% of the students think that a solute 'disappears' when dissolved. These studies indicate that even secondary students struggle with the notion of the conservation of substances, or mass. Another commonly reported confusion concerns dissolving and melting, where both terms are used synonymously or interchangeably, for example: "The sugar is dissolving... the water is sort of melting the sugar crystals" [38] (p. 18). This confusion decreases with age but is still quite common among secondary students. Studies in this field tend to rely on student verbal responses in interviews, sometimes combined with drawings, to identify student conceptual understandings. Therefore, these studies involve a limited number of semiotic modes to generate insight into student reasoning related to the process of dissolving.

4. A Multi-Layered Framework for Multimodal Discourse Analysis

The data was analyzed using a multi-layered framework (see Figure 3), which was first developed in an earlier study that focused on student reasoning in science [12] and refined in this study. This framework allowed the researchers to work with the data across three timescales: macro, meso and micro, to infer how phenomenon at the macro scale was caused by actions and activities at the meso and micro scales [39,40]. We began the analytical process at the macro level by identifying variations in the 'products' of reasoning in the form of student generated artefacts and the repeated viewing of video records of the lesson to acquire a holistic picture of how the lesson unfolded, which led to the identification of interesting segments for further analysis. At the meso level, we coded and transcribed the selected segments of interactions to narrow the analytical focus of the study. At the micro level, we conducted a frame-by-frame analysis to unpack student use of modes in reasoning about dissolving in fine-grain details. In the following sections, we demonstrate how each layer of the analysis could provide useful insight into student reasoning during the science session and the interplays between the semiotic modes in their reasoning.



Figure 3. A multi-layered framework for multimodal discourse analysis.

4.1. Analysis of Student Generated Artefacts

First, the artefacts of student work created during the session were collected and compared. This involved worksheets, photographs of whiteboard drawings and photographs of playdough models. These artefacts provided us with insight into the 'product' of the student reasoning process. Comparing and contrasting student work demonstrated some variations in student representations and generated questions to be further investigated by looking into the video recorded classroom activities.

In this section, we provide an overview of the drawings made by each of the 12 student pairs on a whiteboard, and the way they use the playdough and toothpicks in response to the question "what they might be able to see" if they could use a magnifying glass that "can be zoomed in millions of times", before and after adding icing sugar, and after shaking and waiting for a while (see Figure 2). It should be noted that the students were not taught about particle models prior to participating in this session. The objective was to focus on the way students represent what they thought was happening, based on their prior knowledge. Student-generated representations were classified into categories by applying the evaluation framework of diagrams developed by McLure, Won and Treagust [41].

Although we were not able to collect artefacts from all 12 student pairs (e.g., some groups wiped out their drawings before we could capture them), all whiteboard drawings referred to the phenomena as they were observed, as each pair drew a bottle, the balloon and indicated the liquid level. The suggestion to draw what they might see with a very strong magnifying glass was not responded to by any of these pairs. The use of playdough and toothpicks included a range of representations. Some pairs used the playdough to recreate a physical copy of the bottle or the balloon (1-1, 6-2) without providing an explanation for the phenomena observed (non-explanation). Others referred to particles. Three pairs (3-1, 3-2, 4-1) devised almost identical models, where playdough balls were connected with toothpicks to represent 'particles' or 'molecules'. However, these models were not explicitly related to the dissolving process, or to the solution that was produced as a result of dissolving (mixed description). It is possible that these students copied each other since they were located at the same (3-1 and 3-2) or a neighbouring table (4-1). Two pairs (5-1, 6-1) used the playdough to represent the liquid and the toothpicks to represent the icing sugar (6-1) or the dissolving process (5-1). By pushing the toothpicks into the playdough, the students of 6-1 aimed to represent the sugar's disappearance (macroscopic description). Finally, one pair (2-2) used the playdough to explain the dissolving process by modelling the icing sugar as balls that gradually became smaller and then invisible (mixed description).

4.2. Repeated Viewing and Selecting Student Groups for Further Analysis

The next stage involved repeated watching of the entire session and the identification of instances of interest to the researchers. The initial aim of the analysis was to identify the actions of the students as they first conducted, then reflected and finally described the phenomena observed in the dissolving investigation. The initial viewing of the lesson involved identifying the various modes employed by the students during the science lesson to represent the phenomena, including: (1) Talking; (2) Manipulating; (3) Drawing; (4) Gesturing. This analysis was facilitated by Studiocode, a computer software which allows for a large amount of data to be analyzed in a quick and reliable way. Studiocode allowed the research team to take the large amount of raw data: 6 tables with 2 groups of 2 students on each table (24 students in total) and to quickly identify points of interest with the groups that could be analyzed in detail.

During this stage, our attention was focused on selecting student pairs for further and a more in-depth analysis. The most important criterion for selection was the way students communicated about their different representations of the dissolving phenomena. Specifically, pairs that used talk, gestures, models and drawings to communicate with each other (explaining, discussing, convincing) about the phenomena were selected rather than pairs where students worked more individually. Additional criteria included the quality of the video and audio recording and of the pictures of the artefacts created by the pair. This resulted in the selection of two pairs for a focused analysis.

4.3. Focused Analysis of Selected Video Segments through Coding

After the initial videos were viewed, we were able to identify student pairs of interest whose process could then be examined in more detail, with more complex codes. To perform the analysis, codes were chosen to examine student reasoning at a more indepth level. This phase of analysis focused on the stages of argumentation the students presented in developing their representation. The argumentation codes were developed based on the work by Erduran, Simon and Osborne [42] who built on Toulmin's Argument pattern/model [43]. We list the argumentation codes, definitions, and examples to illustrate the codes from the data, as follows:

- Generating claims: an initial attempt to predict an outcome, e.g., It is going to blow up the balloon.
- Refining claims: providing additional detail to claims already made, e.g., "You know how in the other one it made little ones. This one it's going to make bigger ones".
- Revising claims: Making changes to claims based on additional information, e.g., "I think the balloon will just stay there but then next one it will fizz up with the baking soda".
- Analyzing and interpreting evidence: making sense of observed results from investigations, e.g., "Oh yeah if you swish the water, it barely moves and doesn't come apart".
- Justifying claims: Using evidence to support claims, e.g., "Yeah see it makes like little clumps of it".
- Coordinating explanations: sharing ideas to form one conclusion, e.g., Student 1: "it's all gone." Student 2: "It's dissolved".
- Reaching consensus: agreeing on a final conclusion and recording as evidence, e.g., "Okay now it's dissolved. It's literally dissolved. Write down it's completely dissolved".

However, as the research team tried to use these codes to understand the argumentation that occurred during the lesson it was difficult to reach a clear consensus as to which stage of argumentation the students had reached. It also proved problematic in that the codes were too fine-grained which made it challenging to make sense of the patterns identified from the analysis. Therefore, the videos of the selected student pairs were re-analyzed from an alternative perspective, that is, the phases of inquiry-based learning [44], using the following set of codes:

- Making predictions, e.g., "It is going to blow up the balloon".
- Making observations, e.g., "It is getting clearer".
- Conducting investigations, e.g., "Okay, you open and smell it".
- Communicating findings, e.g., "Perhaps draw a bottle and at the bottom it has little dots".
- Constructing representations/models, e.g., constructing a model to demonstrate dissolving using playdough or drawing a picture.
- Drawing inferences, e.g., "Because the chemicals in the icing sugar are different from the baking soda".

This data was again coded using Studiocode and comparisons were made between codes for the stage of inquiry and the stage of argumentation. Sections of the video, in which codes for the stage of inquiry co-occurred with codes for the stage of argumentation, were selected for transcribing. An example of this analysis is presented in Figure 4.

In this example, a pair of students' actions were coded using the following two sets of codes: argumentation (green) and phases of scientific inquiry (blue). As demonstrated in Figure 4, the scientific inquiry codes seem to focus on a longer timescale compared with the argumentation codes. Argumentation codes identify aspects of the interactions that are related to student claim making during the investigation. This coding relies heavily on verbal, and in particular, the spoken mode of the interactions captured on video. It

identifies 'utterances' as the main unit of analysis. However, the challenge of this coding lies in making inferences about the intention of participants' utterances so that the process of claim making can be meaningfully separated and interpreted. For example, one of the challenges encountered by the researchers was to differentiate between revising claims and refining claims in interpreting student utterances. The grain size of such events is quite small, often involving seconds of speech made by an individual and involve 'high' inference from the coder. In such cases, differentiation becomes challenging if not problematic. In contrast, the scientific inquiry analysis phase draws the researchers' attention to 'chunks' of actions (both verbal and non-verbal) that align well with the unfolding of the meaningmaking events during the student investigations. It requires a relatively 'low' inference in the coding process.

•	•	Mix ar	nd Myles Coded New	
Movi	e Edit New row New merged row	New timeline Text Labels Matrix	Report Organizer Sorter Transcription Database View my codes	Q 👙 Find Settings
_	00:00:00	0 00.04.00.00 00.0800.00 00.12.00.00 00.16.00.00 00.20.00.00 00.24	00.00 00.28.00.00 00.28.00.00 00.36.00.00 00.40.00.00 00.44.00.00 00.48.00.00 00.52.00.00	00:56:00.00 01:00:00.00
1	Generating Claims Analysing and Interpreting Evidence			
3	Revising Claim			
4	Justifying Claims			11/2 (12)
6	Coordinating Explanations Making predictions			
8	Making observations			
10	Communicating findings			20000
11 12	Constructing representations/models Drawing inferences		2/////////////////////////////////////	6

Figure 4. Identification of key segments for further analysis.

4.4. Constructing Intermediate Research Artefacts: Multimodal Transcripts

Next, the transcribed data was linked to the images of representations that the students produced (both drawn on the whiteboard or built using playdough), the written text on the worksheet as well as still shots of the video footage to show the gestures used by the students to develop their reasoning. This set of data was compiled into a document that included time and transcript of the interactions, connected to the representations created by the students and/or still shots of gestures (see Figure 5 below for an example).



Figure 5. An example of an intermediate multimodal transcript.

The research team had an extensive discussions regarding how to display the transcripts so that they incorporate both the verbal and visual aspects of the classroom interactions captured on video. One of the challenges is to display actions that take place simultaneously, for example, speech accompanied by actions. While speech can often be organized in a sequential manner, questions were raised about how best to insert actions into the transcript, to help the reader to make sense of the unfolding of a meaning-making event. Figure 5 is an example of an intermediate research artefact created to support the research team to make understand the way the multimodal phenomena of developing an explanation for dissolving unfolded over time, and the interplay of semiotic modes in supporting students' attempts to develop such an explanation. In the intermediate transcript displayed above, we used brackets to indicate the simultaneity between speech and action. Different colors were used to highlight student spoken words in reference to particles (in green) and the process of dissolving (in yellow).

4.5. Unpacking 'Interesting' Moments of Student Reasoning through Fine-Grained Analysis

The final stage of the analysis was to study, in frame by frame detail, the interaction between pairs of students during the selected fragments. Moments of interest occurred when the students engaged in exchanges about their interpretations of the observed dissolving phenomenon. In order to capture the process in its full complexity, each verbal utterance was linked directly to the student artefacts and to their gestures, using snapshots of video recordings. This frame by frame break-down provided a highly detailed picture of the reasoning of the students as they developed their representations.

To allow for a deeper understanding of how students used a range of modes (speech, gesture, drawing and materials) to represent their ideas and support their reasoning, we proceed by analyzing the interactions in one of the student pairs (i.e., 2-2), who worked side by side until one of the students started to explain his representation to the other. This pair was selected because the students employed a combination of semiotic modes in their interactions.

The process exemplifies how different semiotic modes were used in synchronous ways to enable and support communication between the two students, who aimed to understand the process of dissolving. In the table above, attempt to capture the use of all modes in a concise way to represent the dynamics of the process, which has a duration of 105 s from start to end, by dividing it into 20 steps, and by referring to the time that elapsed between each pair of steps (second column). Each step is then represented by what was said by the student (third column), what the student physically did during this step (e.g., drawing, gazing, gesturing, manipulating playdough; fourth column) and a still shot from the video that represents their actions during this step (fifth column).

The first five steps of this sequence focus on Michael, who was working and talking to himself while breaking up a lump of playdough using a toothpick in a number of iterations into increasingly smaller and smaller bits, until he concluded that there was 'nothing' (Line 5). Next, he turned to his partner, Matthew, who concentrated on drawing on the whiteboard until this point (6). Michael began to repeat his explanation, while pointing at the clumps of playdough. Matthew started to engage with the playdough (10), until Michael concluded his explanation, by asking 'Do you get what I mean?', while opening his hand to support his question (13). Matthew nodded (14), and then Michael continued to say, 'I don't know if that makes sense' (15). Matthew responded, 'Yeah it does' (16) and Michael went on to repeat the breaking-up process, while saying 'it went from that to nothing' (18-19). The excerpt is concluded by Matthew's affirmation (20).

This fine-grained transcript of student speech and actions, while interacting with materials, provides insight into the interactions of different modes to give rise to semantic expansion as the meaning potential of different semiotic modes were integrated [23]. Some of the modes, occurring simultaneously, seem to overlap in conveying the same meaning, for example, Line 14, where the speech 'yeah' expressed agreement is similar to the 'nodding'. In other cases, the speech and the action complemented each other in

communicating ideas. For example, Michael used a number of indexical words, such as 'this', 'that', or 'it', which identified aspects of his playdough model.

From a scientific point of view, the reasoning of the students is interesting since it moves from a macroscopic perspective (i.e., the playdough representing a visible amount of sugar) to increasingly smaller, however still visible clumps, to a level where 'nothing' is left. From our data, we are unable to deduce whether the students' understanding of 'nothing' refers to 'not -or no longer- visible with the naked eye' or 'completely gone'. The latter notion would imply that the students believe that sugar 'disappears' in the dissolving process, in contrast with the notion of conservation of matter (cf., [32]). It must be noted that Michael initially used the term 'particles' (Line 1), but later talked in terms of 'clumps'. These students, unlike some of their classmates, did not use the term 'molecule'.

5. Discussion

5.1. Student Understanding of Dissolving

Our multi-layered analysis has generated insights into student understandings of the phenomenon of dissolving. Similar to what has been reported in the literature, the majority of student representations are of macroscopic or observable levels (Table 1). This is not surprising given that the students were not instructed about explanations of particles prior to working on the tasks. Nevertheless, four of the 12 student pairs incorporated particle ideas in their representations, using playdough rather than drawings. The discussions in the student pairs provided further insight into the reasoning that students applied in comprehending the dissolving process. The use of different modes enabled and supported students in representing their reasoning in a dynamic manner. Previous studies in this area relied on interviews, sometimes in combination with drawings. Such methodologies are limited in generating insights into the process of student reasoning which, as the present study shows, often involves dynamic interplays of multiple modes at a single point in time, especially when student reasoning is focused on processes, such as dissolving, rather than static phenomena, such as a solution. While previous research [37] reported that almost half of the students (age 11-14) in their sample thought that a solute 'disappears', the excerpt from Michael and Matthew reveals their reasoning underlying the process of disappearing.

McLure, Won, and Tregust [41] conducted a study for the same age group (Grade 5 and 6). However, their group were comprised of high achieving students, who were instructed about the particle theory of matter. Data collection was limited to diagrams and verbal statements (probed by group discussion or Socratic questioning by teacher/researcher). They found that some students produced diagrams that were categorized as simple or complex scientific explanations. Our students were not instructed but were instead given an open-ended task that allowed them to generate their representations based on their observations and imagination, without being restricted to verbal modes, leading to multimodal expression. Researchers did not intervene in the process. Unsurprisingly, our students' authentic, or spontaneous, reasoning. The combination of the tasks, the resources we provided to the students during their investigation and the ways in which the data were collected, allowed for the identification of in-depth explanations, evident in this analysis.

Table 1. Student constructed representations of dissolving.

Category	Characteristics of Representation	Table Groups	Sample Representation	Interpretation
Non-explanation (descriptive without providing an explanation for the phenomenon)	Drawings and playdough used to represent equipment (e.g., bottles and balloon) but do not contain explanation for the phenomenon.	1-1, 1-2, 4-2, 6-2		Drawing depicts what was observed.

Category	Characteristics of Representation	Table Groups	Sample Representation	Interpretation
Macroscopic description (focusing on the observable features of the phenomenon)	Drawings or playdough used to represent the process of dissolving.	5-1, 6-1	B	Playdough was used to represent the liquid, toothpicks used to represent the icing sugar. Toothpicks were pushed into the playdough to represent dissolving.
Mixed description (a mixture of particle and observable descriptions of the phenomenon)	Using playdough to construct particles; related or unrelated to observations.	3-1, 3-2, 4-1 2-2		Playdough used to make balls which were joined by toothpicks.
Unable to classify	Constructed representations or models are vague or unrelated to phenomena.	2-1, 5-2		No explanation given for representation.

Table 1. Cont.

5.2. Affordances of the Multi-Layered Analytical Framework for Identifying Student Reasoning

For a long time, technologies constrained how researchers were able to capture and interpret meaning making. Researchers tended to rely on textual and audio records of human interactions in their analysis. The recent rise of video technology and analytical tools has allowed researchers to record, store, and analyse multimodal phenomena and events. However, as Lemke [18] described, "we cannot understand the epistemology of video as representation unless we also understand the process by which we make meaning with video when we experience it" (p. 40). Our multi-layered framework of data analysis (Figure 3) provides an entry point for methodological discussions of analyzing classroom video data.

This multi-layered approach to analyzing multimodal classroom data and exploring student reasoning, allowed us to work with the data at three timescales: macro, meso and micro, to observe how a phenomenon at the macro scale was contributed to by actions and activities at the meso and micro scales [39,40]. Each layer of the analysis provided useful insights into the reasoning of the students during their science investigation and the interplays between the semiotic modes in this reasoning. The analysis of student-generated artefacts identified the 'product' of the reasoning which led to questions for further exploration. The repeated viewings of the video recorded session allowed for the identification of interesting interactions for further analysis. Coding, using different frameworks (phases of inquiry and argumentation), allowed the researchers to zoom in and out of the selected segments of interactions to identify suitable units of analysis and to sharpen the analytical focus of the study. Finally, the frame-by-frame analysis in Table 2 allowed us to 'slow down' the actions captured on video so that we could analyse student multimodal interactions in fine-grain detail.

However, we had to omit certain information. We are limited by the data sources we used and by the format of the written paper. Most importantly, we are unable to report moving images and sound. Thus, we lose information about how students moved their hands and bodies and how they used their voices, for instance, to highlight or emphasise certain observations.

Line	Student	Verbal Utterance	Action	Video Snapshot
1	Michael	(talking out loud) I'm going to make those little particles.	Michael cuts up playdough into small pieces with toothpick.	
2	Michael	So it goes from that	Looking at table, Michael points to his small clump of playdough.	
3	Michael	(talking out loud) to that	Michael continues to break up playdough into smaller clumps with toothpick.	
4	Michael		Michael continues to break up playdough into smaller clumps with toothpick.	
5	Michael	(talking out loud)—and then to nothing.	Michael points at playdough.	
6	Michael	See. Oi look. Hey Matthew.		
7	Michael	See look, you know what I mean?	Michael looks to Matthew and points at his playdough representation.	

 Table 2. Analysis of meaning making during observations of sugar dissolving.

Table 2. Cont.

Line	Student	Verbal Utterance	Action	Video Snapshot
8	Michael	Do you get this? You get what I mean?	Michael looks at Matthew and runs his finger from top to bottom showing the sequence of biggest to smallest.	
9	Michael	It went from that kind of clumps,	Michael points to the largest playdough ball (particle)	
10	Michael	to that kind of clumps	Michael moves finger down to next particle (ball) Matthew begins pointing to the particles (balls) and follows Michael as he moves down from biggest to smallest.	
11	Michael	to that kind of clump,	Michael moves finger down to next particle (ball).	
12	Michael	to that kind of clump.	Michael moves finger down to next particle (ball) and looks at Matthew.	
13	Michael	Do you get what I mean?	Michael looks at Matthew and opens palm.	

14MatthewYeah.Matthew looks at playdough representation and nods.Section and playdough representation and nods.15MichaelYeah. I don't know if that makes sense.Both students look back to the representation.Section and section and pointing to empty16MatthewYeah it does.Section and representation and pointing to emptySection and pointing to empty17MatthewAnd then it went to that nothing.Matthew looks at representation and pointing to emptySection and pointing to empty18MichaelI'm going to try and that nothing.Michael scrapes at the smallest of the particlesSection and pointing to empty space on table.19MichaelTo nothingMichael scrapes the final particles until they are too small.Section and pointing to empty20MatthewYep.	Line	Student	Verbal Utterance	Action	Video Snapshot
15MichaelYeah. I don't know if that makes sense.Both students look back to the representation.SenseSense16MatthewYeah it does.SenseSenseSense17MatthewYeah it does.SenseSenseSense18MatthewAnd then it went to that nothing.Matthew looks at representation and pointing to empty space on table.SenseSense18MichaelClear that out. It went from thatMichael scrapes at the shallest of the particles.Sense19MichaelTo nothingMichael scrapes the shallest or small.SenseSense20MatthewYep.KenterSenseSense	14	Matthew	Yeah.	Matthew looks at playdough representation and nods.	
16MatthewYeah it does.Image: Constraint of the particles of the p	15	Michael	Yeah. I don't know if that makes sense.	Both students look back to the representation.	
17MatthewAnd then it went to that nothing.Matthew looks at representation and pointing to empty space on table.Image: Constraint of the section and pointing to empty space on table.Image: Constraint of the section and pointing to empty space on table.Image: Constraint of the section and pointing to empty space on table.Image: Constraint of the section and pointing to empty space on table.Image: Constraint of the section and 	16	Matthew	Yeah it does.		
18MichaelI'm going to try and clear that out. It went from thatMichael scrapes at the smallest of the particles (balls) with a toothpick.Image: Constant of the particles souther that the particles (balls) with a toothpick.19MichaelTo nothingMichael scrapes the final particles until they are too small.Image: Constant of the particles to small.20MatthewYep.	17	Matthew	And then it went to that nothing.	Matthew looks at representation and pointing to empty space on table.	
19MichaelTo nothingMichael scrapes the final particles until they are too small.Image: Constant of the second	18	Michael	I'm going to try and clear that out. It went from that	Michael scrapes at the smallest of the particles (balls) with a toothpick.	
20 Matthew Yep.	19	Michael	To nothing	Michael scrapes the final particles until they are too small.	
	20	Matthew	Yep.		

Table 2. Cont.

5.3. Research Process as Multimodal Representation Construction

Social semiotics has been used as a methodology to understand how people make meaning using modes available to them. This perspective is also applicable to research processes where the researchers utilize available semiotic resources to construct meanings from multimodal data such as video data of classrooms. In our work, we were interested in how social semiotics, as a theoretical framework and as a methodology, can help us to better understand student reasoning in science. This, in turn, allowed us to reflect on our research as a multimodal representation construction process itself by which we selected, highlighted, and assembled a variety of semiotic resources and modes in representing and communicating our findings, including the construction of intermediate analytical artefacts such as StudioCode timelines and multimodal transcripts. In each layer of the analysis reported in this paper, we, as researchers, had to make decisions about where to look, what to look for, and how to analyse the events identified. While these decisions, such as selecting and highlighting certain aspects of the data, are often guided by the epistemological positions of the researchers, rarely do researchers systematically reflect on and discuss these methodological decisions in research publications.

As Bezemer and Mavers [45] argued, even the process of transcription itself should be considered as a social meaning-making practice by which the empirical phenomena under investigation are reconstructed. Similarly, Ayaß [46] (p. 508) argued that transcription is 'a constitutive part of the empirical research process' whereby the visual and audio records of social events are often transformed into written text, largely guided by the transcriber's interpretive frames and intentions. In our case, we employed video analysis software (i.e., StudioCode) to support our analytical process by drawing our attention to particular groups of students and segments of interactions recorded on video. In selecting the appropriate unit of analysis for identifying students' reasoning, we utilized two different sets of codes, namely, argumentation and inquiry, drawing on the relevant research literature. This allowed us to compare the two sets of coding approaches to identify the analytical affordance of each approach. While argumentation focused our attention on smaller grain sizes (e.g., an utterance), shorter time scales (e.g., in seconds) and required higher inference from the researcher, the inquiry codes tended to draw our attention to larger grain sizes, which encompass not only 'saying' but also 'doing' of the students in each of the coded instances. This inquiry-focused analysis aligned better with our research purpose of identifying interplays between semiotic modes in student reasoning.

The construction of multimodal transcripts 'slowed down' the interesting moments identified in the StudioCode analysis through the combined use of text and still images to identify and highlight the interplays between different modes (e.g., verbal, visual, and kinaesthetic). The decision about what to include and omit in transcripts, as demonstrated in the two examples (Figure 5 and Table 2), shows how the researchers engaged with the selected segments of recorded events in an incremental process of refinement, through which the transcript increasingly focuses on a particular aspect of interest in response to the issues emerging from the selected segments such as interplays of semiotic modes in student reasoning. The transcript in Figure 5 demonstrates the correspondence between speech and visuals through the use of conventions such as colour coding and brackets, to indicate the simultaneity of speech and action. In comparison, the transcript in Table 2 draws attention to the simultaneous occurrence of multiple modes at any given time, guided by our refined analytical methods.

It should be noted that representing social interactions in the form of timelines and transcripts in research publications involves reconstruction and 'transduction' between modes. The reconstruction of the observed interactions into analytical artefacts, such as timelines of events and multimodal transcripts, can help to generate new and fresh insights. In our study, this process of reconstruction, through transduction, helped us to 'see' the unfolding of student reasoning as a multimodal process

6. Conclusions

In this paper, we applied a multi-layered analytic framework for an exploration of classroom communication as multimodal phenomena that focuses on different units of analysis for the purpose of generating insight into student reasoning in science. Our analysis demonstrated the need to use multiple ways to both capture (video, artefacts, pictures) and analyze students' reasoning as multimodal events, involving the coordination of speech, gesture, image, and action. While this framework was developed specifically for

understanding student reasoning in science, we believe it has the potential to be used as a general framework for guiding video analyses of teaching and learning across a range of classroom situations.

Theoretically, building on the intersections of multimodality and discourse analysis perspectives, this analytical framework supported the researchers in identifying the interplays between the different semiotic modes involved in student reasoning concerning a science phenomenon. This focus on the diversity of modes as resources for meaning making provided strong links to recent studies on multi-representations and representational competence [25,26], which are argued to be critical for developing students' conceptual understanding in science.

Methodologically, this framework raised questions about the process of 'reconstruction' involved in research and the need for systematic reflections on methodological decision making in the research process. It highlights that research itself should be considered as a multimodal representation construction process by which the researcher purposefully selects, highlights, and assembles a variety of semiotic resources and modes. We believe that our approach to the data analysis presents an incremental process of refinement, whereby the generation of intermediate analytical artefacts such as transcripts and timelines facilitates our professional vision [47], rendering visible the socially and culturally shaped categories through which the researchers see and reconstruct the world [45]. However, we recognize that our paper is only a modest theoretical and empirical contribution to this fast-growing field of multimodality and discourse analysis in science education. Further research is needed to explore how this proposed framework can be applied and adapted in investigations of student reasoning across a range of science topics and in a range of social and cultural settings.

The present study highlighted the multimodal nature of ideas that students have about science phenomena such as dissolving and therefore, the need to explore the interplays of different modes in student reasoning about such phenomena. The findings of this study can support teachers in understanding the reasoning process as students grapple with the complexity of physical phenomena. Understanding the dynamic interplays between different modes can support science teachers to make decision about how to select and sequence such modes in ways that support and contribute to student learning. However, we acknowledge that the methodology employed in the paper only presents a first step in this direction and requires further refinement to make such dynamic interplays of modes optimally accessible for systematic analysis. The interdependence of modes in generating meaning poses a major challenge for research, to understand the contribution of each mode. Furthermore, how to 'slice' the multimodal phenomenon of interest to identify the varieties of meaning-making at play presents yet another challenge [48]. Our analytical framework offers one method through which to identify these complexities. Finally, this multi-layered framework is both time consuming and labour intensive. Future research in this area is needed to translate such research-informed frameworks and findings into forms that teachers can readily use in assessing student reasoning in their classrooms.

Author Contributions: L.X. and J.v.D. conceptualization and data collection. L.X., J.v.D. and R.H. data analysis and writing, reviewing and editing; L.X.: project administration and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Australian Research Council, grant number SR120300015.

Institutional Review Board Statement: The study was conducted according to the guidelines of The National Statement on Ethical Conduct in Human Research (2007, updated 2018), and approved by the Institutional Review Board (or Ethics Committee) of Deakin University (Reference number. HAE-17-029 approved on 9 June 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethics requirements of maintaining data privacy for the reported study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Gee, J.P.; Green, J.L. Discourse Analysis, Learning, and Social Practice: A methodological study. *Rev. Rev. Educ.* 1998, 23, 119–169. [CrossRef]
- 2. Erickson, F. A history of qualitative inquiry in social and educational research. In *The Sage Handbook of Qualitative Research;* Lincoln, Y.S., Denzin, N.K., Eds.; Sage Publishing: Thousand Oaks, CA, USA, 2011; Volume 4, pp. 43–59.
- 3. Scott, P.; Mortimer, E. Meaning making in high school science classrooms: A framework for analysing meaning making interactions. In *Research and the Quality of Science Education*; Boersma, K., Goedhart, M., De Jong, O., Eijkelhof, H., Eds.; Springer: Dordrecht, Germany, 2005; pp. 395–406.
- 4. Lemke, J.L. Talking Science: Language, Learning, and Values; Ablex Publishing Corporation: Norwood, NJ, USA, 1990.
- 5. Yore, L.D.; Hand, B. Epilogue: Plotting a research agenda for multiple representations, multiple modality, and multimodal representational competency. *Res. Sci. Educ.* **2010**, *40*, 93–101. [CrossRef]
- 6. Kress, G.; Jewitt, C.; Ogborn, J.; Tsatsarelis, C. *Multimodal Teaching and Learning: The Rhetorics of the Science Classroom*; Continuum: London, UK; New York, NY, USA, 2001.
- Moro, L.; Mortimer, E.F.; Tiberghien, A. The use of social semiotic multimodality and joint action theory to describe teaching practices: Two cases studies with experienced teachers. *Classr. Discourse* 2020, *11*, 229–251. [CrossRef]
- 8. Tytler, R.; Prain, V.; Hubber, P.; Waldrip, B. *Constructing Representations to Learn in Science*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
- 9. Tang, K.-S. The interplay of representations and patterns of classroom discourse in science teaching sequences. *Int. J. Sci. Educ.* **2016**, *38*, 2069–2095. [CrossRef]
- 10. Tang, K.-S.; Won, M.; Treagust, D. Analytical framework for student-generated drawings. *Int. J. Sci. Educ.* 2019, 41, 2296–2322. [CrossRef]
- 11. Tsatsarelis, C.; Ogborn, J.; Jewitt, C.; Kress, G. Rhetorical construction of cells in science and in a science classroom. *Res. Sci. Educ.* **2000**, *30*, 451–463. [CrossRef]
- 12. Xu, L.; Ferguson, J.; Tytler, R. Student reasoning about the lever principle through multimodal representations: A socio-semiotic approach. *Int. J. Sci. Math. Educ.* 2021, *19*, 1167–1186. [CrossRef]
- 13. Erickson, F. Origins: A brief intellectual and technological history of the emergence of multimodal discourse analysis. In *Discourse and Technology: Multimodal Discourse Analysis;* Levine, P., Scollon, R., Eds.; Georgetown University Press: Washington, DC, USA, 2004; pp. 196–207.
- 14. Goldman, R.; Pea, R.; Barron, B.; Derry, S.J. Video Research in the Learning Sciences; Routledge: New York, NY, USA; London, UK, 2014.
- 15. Harris, A.M. Video as Method; Oxford University Press: Oxford, UK, 2016.
- 16. Jewitt, C. An introduction to using video for research. 2012, unpublished work. Available online: Eprints.ncrm.ac.uk/2259/ (accessed on 1 August 2021).
- 17. Jordan, B.; Henderson, A. Interaction Analysis: Foundations and Practice. J. Learn. Sci. 1995, 4, 39–103. [CrossRef]
- 18. Lemke, J. Video epistemology in-and-outside the box: Traversing attentional spaces. In *Video Research in the Learning Sciences;* Goldman, R., Barron, B., Pea, R., Derry, S., Eds.; Routledge: New York, NY, USA; London, UK, 2007; pp. 39–51.
- 19. Norris, S. Analyzing Multimodal Interaction: A Methodological Framework; Routledge: London, UK, 2004.
- 20. Jewitt, C. Multimodality and literacy in school classrooms. Rev. Res. Educ. 2008, 32, 241-267. [CrossRef]
- 21. Halliday, M.A.K.; Martin, J.R. Writing Science: Literacy and Discursive Power; Routledge: New York, NY, USA; London, UK, 2003.
- 22. Lemke, J.L. Teaching all the languages of science: Words, symbols, images, and actions. In Proceedings of the Conference on Science Education, Barcelona, Spain, 1998. Available online: http://academic.brooklyn.cuny.edu/education/jlemke/papers/barcelon.htm (accessed on 1 August 2021).
- 23. O'Halloran, K.L. Multimodal discourse analysis. In *Companion to Discourse*; Hyland, K., Paltridge, B., Eds.; Continuum: London, UK; New York, NY, USA, 2011; pp. 120–137.
- 24. Jewitt, C.; Bezemer, J.; O'Halloran, K. Introducing Multimodality; Routledge: New York, NY, USA; London, UK, 2016.
- 25. Prain, V.; Tytler, R.; Peterson, S. Multiple representation in learning about evaporation. *Int. J. Sci. Educ.* 2009, *31*, 787–808. [CrossRef]
- 26. Waldrip, B.; Prain, V. Engaging students in learning science through promoting creative reasoning. *Int. J. Sci. Educ.* **2017**, *39*, 2052–2072. [CrossRef]
- 27. Airey, J.; Linder, C. A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *J. Res. Sci. Teach.* 2009, *46*, 27–49. [CrossRef]

- Tang, K.-S.; Delgado, C.; Moje, E.B. An integrative framework for the analysis of multiple and multimodal representations for meaning-making in science education. *Sci. Educ.* 2014, *98*, 305–326. [CrossRef]
- 29. Jaipal, K. Meaning making through multiple modalities in a biology classroom: A multimodal semiotics discourse analysis. *Sci. Educ.* **2010**, *94*, 48–72. [CrossRef]
- 30. Ainsworth, S. DeFT: A conceptual framework for considering learning with multiple representations. *Learn. Instr.* 2006, 16, 183–198. [CrossRef]
- 31. Márquez, C.; Izquierdo, M.; Espinet, M. Multimodal science teachers' discourse in modeling the water cycle. *Sci. Educ.* **2006**, *90*, 202–226. [CrossRef]
- Tversky, B. Spatial Schemas in Depictions. In Spatial Schemas and Abstract Thought; The MIT Press: Cambridge, MA, USA, 2001; pp. 79–111.
- 33. Tang, K.-S.; Tan, S.C.; Yeo, J. Students' multimodal construction of the work–energy concept. *Int. J. Sci. Educ.* **2011**, *33*, 1775–1804. [CrossRef]
- 34. White, R.; Gunstone, R. Probing Understanding; Routledge: New York, NY, USA; London, UK, 2014.
- 35. Piaget, J.; Inhelder, B. *The Child's Construction of Quantities: Conservation and Atomism*; Routledge: New York, NY, USA; London, UK, 1974; Volume 2.
- 36. Driver, R.; Guesne, E.; Tiberghien, A. Children's Ideas in Science; Open University Press: Milton Keynes, UK, 1985.
- 37. Prieto, T.; Blanco, A.; Rodriguez, A. The ideas of 11 to 14-year-old students about the nature of solutions. *Int. J. Sci. Educ.* **1989**, *11*, 451–463. [CrossRef]
- Cosgrove, M.; Osborne, R. *Physical change. Learning in Science Project;* Working Paper No. 210; University of Waikato: Hamilton, New Zealand, 1985; pp. 28–57.
- Lemke, J.L. Across the Scales of Time: Artifacts, activities, and meaning in ecosocial systems. *Mind Cult. Act.* 2000, 7, 273–290. [CrossRef]
- 40. Tang, K.-S. Discourse Strategies for Science Teaching and Learning: Research and Practice; Routledge: New York, NY, USA; London, UK, 2020.
- 41. McLure, F.; Won, M.; Treagust, D.F. Analysis of Students' Diagrams Explaining Scientific Phenomena. *Res. Sci. Educ.* 2021, 1–17. [CrossRef]
- 42. Erduran, S.; Simon, S.; Osborne, J. Tapping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Sci. Educ.* **2004**, *88*, 915–933. [CrossRef]
- 43. Toulmin, S.E. The Uses of Argument; Cambridge University Press: Cambridge, UK, 1958.
- 44. Pedaste, M.; Mäeots, M.; Siiman, L.A.; De Jong, T.; Van Riesen, S.A.; Kamp, E.T.; Manoli, C.C.; Zacharia, Z.C.; Tsourlidaki, E. Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educ. Res. Rev.* **2015**, *14*, 47–61. [CrossRef]
- Bezemer, J.; Mavers, D. Multimodal transcription as academic practice: A social semiotic perspective. *Int. J. Soc. Res. Methodol.* 2011, 14, 191–206. [CrossRef]
- 46. Ayaß, R. Doing data: The status of transcripts in Conversation Analysis. Discourse Stud. 2015, 17, 505–528. [CrossRef]
- 47. Goodwin, C. Professional Vision. Am. Anthropol. 1994, 96, 606–633. [CrossRef]
- 48. Bateman, J.; Wildfeuer, J.; Hiippala, T. Multimodality: Foundations, Research, and Analysis, A Problem-Oriented Introduction; Walter de Gruyter GmbH: Berlin, Germany, 2017.