THE NATURE OF SCIENTIFIC EXPLANATION (NOSE): USING A PHILOSOPHICALLY GUIDED FRAMEWORK TO EXAMINE THE NATURE AND QUALITY OF SCIENTIFIC EXPLANATIONS CONSTRUCTED BY FRESHMAN COLLEGE STUDENTS, SCIENCE TEACHERS, AND PRACTICING SCIENTISTS

BY

SAHAR ALAMEH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Curriculum and Instruction
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Professor Emeritus David Brown, Chair
Professor Fouad Abd-El-Khalick, University of North Carolina at Chapel Hill, Director of Research
Associate Professor Gloriana González Rivera
Professor Saouma BouJaoude, American University of Beirut
Teaching Associate Professor Barbara Hug

ABSTRACT

Issues regarding scientific explanation have been of interest to philosophers from Pre-Socratic times. The notion of scientific explanation is of interest not only to philosophers, but also to science educators as is clearly evident in the emphasis given to K-12 students' construction of explanations in current national science education reform efforts – the Next Generation Science Standards NGSS (NGSS Lead States, 2013). Nonetheless, there is a dearth of research on conceptualizing explanation in science education. Scientific explanation seems to be ill-defined (or left undefined) among researchers, science teachers and, in turn, students (Braaten & Windschitl, 2010, p. 639).

Guided by philosophical models of and approaches to explanation, this study proposed a framework – the Nature of Scientific Explanation (NOSE) – for assessing the type, nature and quality of scientific explanations. Furthermore, to establish the validity and usefulness of the NOSE framework, the study aimed to (a) examine college freshman science students', secondary science teachers', and practicing scientists' explanations, (b) elucidate their perceptions of explanations and how they compare to the formal analytical NOSE framework and (c) characterize the nature of the criteria that participant students, teachers, and scientists deploy when assessing the "validity" of explanations. The following research questions guided the study: (1) How do college freshmen science students', secondary science teachers' and practicing scientists' explanations fare when assessed using the NOSE framework? In other words, what is the nature (structural elements) and quality of participants' scientific explanations when analyzed using the NOSE framework? (2) How do college freshmen science students', secondary science teachers' and practicing scientists' explanations of scientific phenomena compare and contrast when analyzed using the NOSE framework? (3) What criteria do college

freshmen science students, secondary science teachers, and practicing scientists use in judging the quality of scientific explanations? How are these criteria consistent among and/or different across the three groups? (4) To what extent are freshmen science students', secondary science teachers' and practicing scientists' views of the quality of scientific explanations aligned with those of NOSE framework?

The study was exploratory in nature. In-depth, semi structured interviews served as the main instrument of data collection. In two separate interviews, participants first constructed explanations of everyday scientific phenomena and then provided feedback on the explanations constructed by other participants. Participants comprised three groups from a large, Midwestern University and neighboring communities: freshman college students, secondary science teachers, and practicing scientists. Each group comprised 10 participants (50% male, 50% female).

The study was conducted in two phases. First, during semi-structured individual interviews all participants generated explanations of various scientific phenomena. Interview transcripts were used to generate an explanation map for each participant following procedures of the NOSE framework developed in this study. During the second phase of the study, participants in each group assessed and provided feedback on the explanations generated during the first phase by other participants. The assignment of explanations to be examined was randomized and ensured that each participant assessed all four scenarios. This examination took place in the context of a second, semi-structured interview. All interviews were audiotaped and transcribed verbatim for analysis.

Data analysis comprised three phases. The first involved (a) the construction of explanation maps from participant transcripts; (b) analysis of maps and corresponding transcripts for emerging participant criteria; (c) using the NOSE framework to generate a profile of participants' types and quality of explanations articulated during the first interview; (d) the explanation maps for each group of participants (students, science teachers, and scientists) were examined to generate a full descriptive account or *profile* of these maps. This analysis resulted in three profiles, one each for the group of participants; and (e) finally the profiles were compared and contrasted to make assertions regarding ways in which students, teachers, and scientists' explanations were similar or different from NOSE framework analysis.

The second phase focused on analyzing transcripts generated during the second interview to characterize participants' perceptions of the nature of explanations, and derive the criteria deployed by members of the three groups to judge the "validity" or "goodness" of explanations. This resulted in individual profiles as to perceptions of the nature of explanations and criteria used to judge explanations. Profiles within each group of participants were analyzed for general patterns to generate a common set of criteria that each group used in their assessment, when applicable. These common sets were then compared and contrasted across the three groups.

The third phase of data analysis focused on comparing and contrasting the sets of criteria derived from the second phase with those NOSE framework. Analysis in this third phase was more conceptual in nature and focused on how the three groups of participants fared in terms of explanation when their explanations were analyzed using NOSE framework.

In general, major findings showed that, when analyzed using NOSE framework, participant scientists did significantly "better" than teachers and students. What is more, most participants across all three groups judged as "best" or "complete" or "good" the explanations made by participant scientists, even though group memberships of the explainers were held anonymous. In addition, scientists had more adequate scientific explanations, from a NOSE perspective, in the sense of providing more relevant and accurate structural elements. Analysis showed that participant explanation maps demonstrated similarities and differences across the three groups. Mainly, scientists' explanations included more pieces of knowledge and lawlike statements, which were relevant and accurate and/or based on prior content knowledge compared to students' and teachers' explanations.

Participants' perceptions of explanations differed significantly. Students tended to think of explanation as a "true" answer to a *why*-question based on observations. However, teachers and scientists tended to perceive explanation as a testable and verifiable tool that provides understanding. More important were the criteria that participants used to assess explanations. Context-dependence and learner-dependence turned out to be two of the most important aspects of explanations considered by participants.

In conclusion, the present study highlights the need articulated by many researchers in science education to understand additional aspects specific to scientific explanation. The study highlighted the importance of not only the structural elements that make up a scientific explanation, but also the connectedness of these elements within the context of teaching and learning. The present findings provide an initial framework for judging the validity of students' and science teachers' scientific explanations.

ACKNOWLEDGMENTS

First and foremost I would like to extend my sincerest gratitude to my dissertation director, Fouad-Abd-El-Khalick for his support and mentorship. His continued guidance and encouragement during the past six years were invaluable. I also want to thank deeply, my advisor, David Brown, for always challenging me to go further and deeper.

I will forever be thankful to my former advisor and current committee member, professor Saouma BouJaoude. My success would not have been possible without Dr. Saouma's support and nurturing. I would also like to extend my thanks and appreciation to my committee members, professors Gloriana González Rivera, and Barbara Hug for their valuable guidance and advice. Many thanks to all the scientists, teachers and students who participated in my study. Special thanks go to Beth Niswander for always being there for me, and for Saad Shehab for his aid in helping me with the reliability of the results.

I want to thank my loving family: my mother, Siham, my brothers, Hassan and Ihab, and my sister, Soha, for their love and support during all the difficult and sad times we passed through. This whole journey would not have been possible without their enormous support and understanding throughout my stay in the United States. To my parents-in-law, April and Henry, I want to thank you for your endless support and dedication.

Finally, and above all, I cannot begin to express my unfailing gratitude and love to my husband, Kent, who has supported me throughout this process and has constantly encouraged me when the tasks seemed arduous and insurmountable.

Dedicated to the loving memory of my father,

Khalil Alameh

He is the reason why I am where I am today.

TABLE OF CONTENTS

CHAPTER 1: THE PROBLEM

Introduction

Issues regarding scientific explanations have been of interest to philosophers from Pre-Socratic times. However, research on explanation in philosophy of science began in earnest with Carl Hempel's (1948) development of the Deductive-Nomological (DN) model of scientific explanation. This model had its supporters (e.g., Gardiner, 1959; Nagel 1961) and detractors (e.g., Hausman, 1998; Salmon, 1989). In spite of its advantages and drawbacks, the DN model is considered to be the milestone of subsequent discussions, and the start of an extensive philosophical research program on scientific explanation.

The concept of scientific explanation is, to a large extent, still vague and ambiguous even in philosophy of science (de Regt, 2009). The reason for this ambiguity, de Regt argues, is the fact that philosophers of science have not reached a consensus regarding the nature of scientific explanation despite extensive debates. In spite of the philosophical debates, researchers in science education can still benefit in significant ways from closely examining and drawing on philosophical accounts of scientific explanation in the form of an adaptive schema of explanations that incorporates some or all of these models into science education. In some science topics, events can be explained by referring to general laws (the Deductive Nomological model), highly probable laws (the Inductive Statistical model), and/or causal mechanistic processes (the Causal Mechanical model) within a pragmatic approach that considers students' levels, their prior knowledge as well the context of learning.

The notion of scientific explanation has been of interest not only to philosophers of science but also to the field of science education – scholars, researchers, teachers, and students. The Next Generation Science Standards (NGSS) (NGSS Lead States, 2013), which embody the current major reform effort in science education in the U.S., list eight essential practices of science and engineering with which all students should engage. "Constructing scientific explanations" is considered one of these essential practices, in addition to "engaging in argument from evidence," "developing and using models," and "obtaining, evaluating, and communicating information" (see also National Research Council [NRC], 2012). Indeed, in the 33-page appendix to the *NGSS* that details the eight practices, the terms "explanation" and "explain" appear 113 times as compared to one single appearance of the verb "argue," none for "argumentation," and 31 instances of terms such as "communicate" and "communication."

In the *NGSS*, scientific explanation is presented in, at least, three different ways: as a goal for science, a tool for learning about science, and a way of answering scientific questions. However, the *NGSS* do not offer any detailed conceptualization of the nature of scientific explanation. The term 'explanation' barely provides the sort of clarity needed to design science instruction, promote students' abilities to construct scientific explanations, and/or assess their progress and mastery of such a central and multilayered practice. For example, the NGSS state that "the goal of science is to construct explanations for the causes of phenomena" (Appendix F, p. 11). While such a statement is important in highlighting the explanatory power of science, it does little to elucidate the meaning of explanation.

The situation is similar in international science curricular documents: constructing explanations is mentioned but its meaning is discussed only peripherally. For example, the National Curriculum of England (NCE) (NCE, 2015) states that students at Key Stage 3 "should be encouraged to relate scientific explanations to phenomena in the world around them and start to use modelling and abstract ideas to develop and evaluate explanations" (p. 59). However, the document does not offer ways to attain such a goal nor does it provide any functional definitions to the term ‘scientific explanation.’ Another example is the Australian Curriculum, Assessment and Reporting Authority (ACARA) (ACARA, 2015), which regards science as a field that seeks “to improve our understanding and explanations of the natural world” and that involves the “construction of explanations based on evidence [and] the development of science concepts, models and theories [which is] dynamic and involves critique and uncertainty” (p. 11). While such curricular goals are important, the ACARA document does not include any information or guidelines for teachers that aim at achieving these objectives. Likewise, in the Spanish Organic Law on the Improvement of the Quality Of Education (LOMCE) (LOMCE, 2015), there is an emphasis on the importance of achieving the ability to “formulate hypotheses to explain everyday phenomena using theories and scientific models” (p. 258), but the document does not include any definitions of explanation or any further elaborations on the ways to explain phenomena as stated. It appears that in science education, as with many other fields, “we are addicted to explanation, constantly asking and answering *why*-questions” (Lipton, 2004, p.1). However, what is lacking in science education is an explication of the nature of explanation.

An examination of studies on scientific explanation in science education indicates that one of the main problems faced by science educators who attempt to assess students’ explanations is that the statements that are analyzed and considered as explanations are sometimes merely answers to questions (e.g., Yang & Wang, 2013). Of those statements that might be explanatory, they are usually examined in unstructured and unsupported manner, such as using Toulmin or a modified version of Toulmin’s model of argumentation (Toulmin, 1958) to assess explanations. The lack of a framework specifically tailored to examining scientific explanation might be the reason for the absence of coherent analysis and evaluation of students’ explanations in the science classroom. This might also explain the well-documented difficulties that teachers and, in turn, students face while constructing explanations (e.g., Erduran, Simon, & Osborne, 2004). In fact, research in science education has shown that teachers do not have adequate understanding of the ways by which explanations are constructed and evaluated; thus they face difficulties in teaching about and assessing their students' explanations (e.g., Haefner & Zembal-Saul, 2004). In addition, Yao, Guo, & Neumann (2016) found that science teachers do not have the appropriate skills and expertise to scaffold students' explanation constructions, while Sadler's (2006) research revealed that some science teachers do not regard constructing explanations as an essential goal in science.

Equally important, apart from a few studies (e.g., Braaten & Windschitl, 2010; Brewer, Chinn, & Samarapungavan, 2000; Kampourakis, Silveira, & Strasser, 2016; Woody, 2013; Yao, et al., 2016) there is a dearth of research on conceptualizing scientific explanation in the field of science education. In particular, scientific explanation seems to be ill-defined (or left undefined) among researchers, science teachers and, in turn, science students. Researchers in science education recently have been calling for a more "clearly articulated conceptualization of scientific explanation for science education" (Braaten & Windschitl, 2010, p. 639).

Statement of the Problem

Despite the ongoing emphasis on explanations in the science classroom, there seems to be no well-articulated frameworks that support students in constructing adequate scientific explanations or that help teachers assess student explanations. The teaching and scaffolding of scientific explanations remain underemphasized in the science classroom (e.g., Zangori, Forbes & Biggers, 2013). Researchers in science education have developed some guidelines for using scientific explanations (e.g. Parnafes, 2012; Metz, 1991; Russ, Scherr, Hammer, & Mikeska, 2008), but these guidelines still have some gaps. Some of the gaps include, among other issues, the lack of clear distinctions between explanations and non-explanations (such as descriptions and predictions), and misinterpretation or misrepresentation of philosophical models of explanation.

An examination of the relevant literature in science education indicates that the ways in which researchers in science education have studied scientific explanation, in most cases, leaves much to be desired in terms of accuracy and completion (Alameh & Abd-El-Khalick, 2018). In fact, research about teaching and learning of explanation in science classrooms must be guided by explicit models or frameworks that specify elements involved in constructing explanations that are particularly applicable to science. More importantly, the development of such models or guidelines should be based on, and consistent with, philosophical scholarship on scientific explanation. In the absence of philosophically-grounded guidelines, research on the teaching and learning and assessment of learners' scientific explanation has often resorted to models that are, at best, peripherally relevant to the topic, such as Toulmin's model of argumentation, without necessarily making a convincing case that arguments are some type of explanation.

Therefore, it is clear that the ways by which researchers in science education have studied the teaching, learning, and assessment of scientific explanation in science classrooms, at best, leaves much to be desired and, at worst, are simply incomplete. It is important for science education researchers to recognize what needs to be addressed when it comes to explanations in science classrooms.

Purpose of the Study

This study aimed to, first, propose a domain-specific framework that is specifically developed for assessing scientific explanation of phenomena in physical science in science classrooms: The Nature of Scientific Explanation (NOSE) framework. This framework is grounded in philosophical models of, and approaches to, explanation. The NOSE framework facilitates the analysis and assessment of students' scientific explanations. For purposes of this study, the framework is intended mainly for use by science education researchers. The framework enables researchers to gain a better understanding of the nature of students' scientific explanations and provides a philosophically-grounded approach to examine and assess whether student-constructed explanations can be considered explanatory or not.

Second, the study aims to elucidate and compare college freshmen science students', secondary science teachers', and practicing scientists' scientific explanations and their views of scientific explanations. This comparative approach follows the work of Abi-El-Mona and Abd-El-Khalick (2011) on perceptions of argumentation among students, teachers, and scientists. In particular, this study aims to: (a) analyze students', teachers', and scientists' scientific explanations using the NOSE framework and determine whether and how NOSE identifies similarities and differences among the three groups; (b) explore how participants' views of explanations fare when examined from the NOSE framework; and (c) elucidate the criteria that participants use in analyzing scientific explanations and compare it with the criteria of NOSE framework. It is worth noting that in this study students, teachers and scientists partook in an interview and provided explanations of various scientific phenomena. Following the first interview, they then participated in a second interview in which they assessed and provided feedback on explanations generated by them as well as by others from the three participating groups during the first interview. The time period between the two interviews was designed to allow time to analyze the data obtained during the first interview, transcribe segments of the audiotape of participants' explanations, and generate the corresponding explanation maps.
The study was guided by the following research questions:

(1) How do college freshmen science students', secondary science teachers' and practicing scientists' explanations fare when assessed using the NOSE framework? In other words, what is the nature (structural elements) and quality of participants' scientific explanations when analyzed using the NOSE framework?

(2) How do college freshmen science students', secondary science teachers' and practicing scientists' explanations of scientific phenomena compare and contrast when analyzed using the NOSE framework?

(3) What criteria do college freshmen science students, secondary science teachers, and practicing scientists use in judging the quality of scientific explanations? How are these criteria consistent among and/or different across the three groups?

(4) To what extent are freshmen science students', secondary science teachers' and practicing scientists' views of the quality of scientific explanations aligned with those of NOSE framework?

Significance of the Study

Evidently, there is a lack of conceptualization of the nature of scientific explanation in science education documents and science curricula (Braaten & Windschitl, 2011), which results in teachers struggling to guide their students to build adequate scientific explanations (e.g., Russ et al., 2008). Thus, the development of NOSE was an effort to fill the existing gap by constructing a framework specific to scientific explanations that is grounded in philosophical models of scientific explanation. In addition, the study shed light on the utility of the NOSE framework. Scientific explanations constructed by college freshmen science students, secondary science teachers, and practicing scientists provided rich data that served to establish the usefulness of the NOSE framework in analyzing scientific explanations. The NOSE framework could also aid in the development of effective instructional interventions that enable students to develop adequate scientific explanations.

Not only did this study aim to propose a framework specific to explanations that is also supported by philosophical models, but it also aimed at providing an empirical support to the usefulness of the developed framework. By doing so, the study helps direct the attention of current research on scientific explanation in science education to the ways that students and science teachers perceive the nature of scientific explanations, and the ways by which practicing scientists judge the adequacy of scientific explanations in the context of science teaching and learning. Furthermore, the proposed study sought to assess whether or not NOSE – a formal analytical framework guided by philosophical models that examines explanations – placed realistic expectations on students' construction and assessment of scientific explanations.

NOSE is among the first attempts in science education to develop a functional framework of scientific explanation guided by the underlying philosophical models that is useful for K-12 science teaching and learning. Synthesizing the applicability of philosophical models into science education is important for science educators to explore students' scientific explanations. It is worth noting that the NOSE framework proposed in this study was not set in stone but was, rather, emergent. It was responsive to the empirical data collected in the study. In particular, the data suggested the need for additional categories to account for certain types of explanations. In addition, the NOSE framework has a summative function (i.e., used to compare learner explanations to canonical explanations), as well as a formative function. The NOSE framework also aims to help guide learners characterize the elements of their explanations and guide them to generate more complete and high quality explanations irrespective of whether these are canonical or not.

CHAPTER 2: REVIEW OF THE LITERATURE

Science educators and major science education organizations are increasingly emphasizing for the importance of students constructing meaningful scientific explanations (e.g. NGSS Lead States, 2013; NRC, 2012). To this date, there has been no clear and articulated definition of the notion of scientific explanation (e.g., Braaten and Windschitl, 2010; Yao, et al., 2016). Nevertheless, there is much work on scientific explanation in philosophy of science. While there is no general agreement on one definition of scientific explanation in philosophy of science, models of explanations provide sound philosophical support (e.g., Achinstein, 1984; Hempel & Oppenheim, 1948; Salmon, 1984).

Knowing *what* happens, although valuable, is not sufficient. Not only do people want to know what happens, they also want to understand *why* it happens. Science can provide answers to why natural phenomena occur. In fact, there is a general agreement that science does not aim to solely describe the world, but mainly to provide "understanding, comprehension, and enlightenment" (Salmon, 1984, p.9). And science attains such goals by providing scientific explanations.

Before delving into definitions of explanation, it is worth noting two important contrasts related to the notion of scientific explanation. The first is a contrast between explanations that are unique to science and those that are not (e.g., explanations related to daily life). The second is a contrast between 'explanations' and 'non-explanations' (such as description, reasoning, interpretation, etc.) within science (Woodward 2003). The contrast between scientific explanation and everyday explanation has been a focus of philosophers and science educators. For example, Brewer et al. (2000) proposed a set of criteria that characterizes explanations in everyday life (empirical accuracy, scope, consistency, simplicity, and plausibility), and concluded that explanations in science are evaluated by the same criteria, plus three others (precision, formalism, and fruitfulness). The current study was concerned solely with scientific explanations—what was considered as scientific explanation and what was not, but a brief summary of the demarcation between scientific and non-scientific explanations is presented in the following section.

Explanation has several uses, many of which were beyond the focus of this study. For example, people usually ask to explain the meaning of a word, the meaning of an anecdote, or the meaning of a metaphor. People also might ask someone to explain to them how to get to a certain place, how to ride a bike, and so on. None of these examples require scientific explanation of a natural phenomenon. Thus, it is important to distinguish between scientific explanations from other types of explanations. In discussing the main differences between scientific and nonscientific explanations, it is important to note that none of the questions in the examples above are asking *why*-questions. In many cases that involve everyday explanations, people ask *what* something means, or *what* is wrong with something. In other cases, they ask *how* to prove something (e.g., mathematically), or *how* to get somewhere. Another type of explanation focuses on how to perform a certain task or an activity (Salmon, 1984).

Explanations of meanings and of how to perform certain tasks are abundant in science. The meaning of a scientific word can be found in a scientific textbook. In addition, a scientist might explain to a mechanic how to construct a dynamo for example. When a scientific explanation is requested, however, one can always ask a *why*-question of some sort (Salmon, 1984). In fact, philosophers have argued that even if the question is not originally formulated as a *why*-question, it can be rewritten as one without changing its meaning (e.g., Bromberger, 1966; van Fraasen, 1980). However, it is important to note that not all *why*-questions call for scientific explanations.

One might ask *why* did an employee get fired – to which the answer can entail a moral or legal justification; or why did someone go to the gas station – a question that requests practical justification (for more examples see Salmon, 1984, pp.9-10).

Another distinction made between scientific explanation and other types of explanation is the idea that scientific explanation aims to simplify or reduce the unfamiliar to the familiar (e.g., Laplace, 1951). For example, Newton's laws explain that comets are objects that behave the same way as planets, whose types of motion are familiar to us. However, the idea of reducing the unfamiliar to the familiar has been refuted by many philosophers: there are plenty of scientific explanations that do not necessarily appeal to everyday experiences (such as Pauli exclusion principle, the mean free path of photons, etc.) but are still explanatory (Salmon, 1984). Another characterization emphasizes the idea that scientific explanation consists in showing that what seems to happen randomly in the world does in fact exhibit some regularity. This characterization does not hold without exceptions: while laws of classical physics include a set of explanations that show natural regularities, quantum physics does not.

Having briefly discussed some of the major differences between scientific and non-scientific explanations, the focus of the NOSE framework is to assess students' scientific explanations within a K-12 context. Thus, an assertion is made that such explanations request, by and large, answers to *why* rather than *what* questions. Moreover, it is important for science education researchers to recognize what needs to be addressed when it comes to explanations in science classrooms. In addition, research about teaching and learning of explanation in science classrooms must be guided by explicit models or frameworks that specify the elements involved in constructing explanations particularly applicable to science. More importantly, the development of such models or guidelines should be based on theoretical and philosophical foundations. In the absence of these guidelines, research on scientific explanation has resorted to models on peripheral topics, such as Toulmin's model of argumentation, to assess explanations without necessarily making a convincing case that arguments are some type of explanation (e.g., Delen and Krajcik 2018; Peker and Wallace 2011; Yang and Wang 2014).

In order to develop these frameworks or guidelines, an outline and a clarification of the models of scientific explanation developed by philosophers of science are needed. The first part of this chapter starts by summarizing the ideas of Carl Hempel and his seminal work on scientific explanation. Then, it presents the problems that other philosophers have raised with Hempel's view of explanation. These problems resulted in the development of new models of explanations, such as causal and causal-mechanical models of scientific explanation (e.g., Cartwright 1983; Salmon 1989), unification models of scientific explanation (e.g., Kitcher 1989), and pragmatic models of scientific explanations (e.g., Achinstein 1984). From the most recent works on scientific explanation, the pragmatic approach to studying scientific explanations developed by Weber, Van Bouwel, and De Vreese (2013) is then presented. This approach suggests a toolbox for analyzing scientists' scientific explanations. The toolbox provides a useful instrument to science education. In this chapter, the summary of these philosophical models is conducted within the context of science education. In particular, examples from science curricula, such as the NGSS, and other explanations related to the science classroom are used to further clarify the philosophical models within a science education context of learning.

In the second part of this chapter, a critical examination of research on scientific explanation in science education is presented. Research on scientific explanation in the science classroom has been of interest to science education researchers for over 40 years. Studies in this regard have addressed various aspects related to scientific explanation construction and tackled different issues associated with the construction of explanation.

PART 1: Philosophical Theories of Scientific Explanation

Hempel's Account of Scientific Explanation

Carl Hempel and Paul Oppenheim (1948) and Hempel (1965) developed an account of scientific explanation that is known as the Covering-law model. This account consists of two models of explanation: The Deductive-Nomological (DN) and the Inductive-Statistical (IS) models of explanation. In his chapter on studies in the logic of explanation, Hempel (1965) considered an explanation of a natural phenomenon as one that answers a *why* rather than a *what* question. For example, an explanation of a natural phenomenon answers the question "why did something happen?" rather than "what happened?"

The Deductive-Nomological (DN) model of explanation. The main purpose of the DN model is to elucidate the necessary and sufficient conditions for scientific explanation. A deductive-nomological explanation includes a deductive composition of statements regarding natural phenomena that are logical consequences of general laws of nature (Hempel, 1965). According to the DN model, a scientific explanation consists of two parts. The first part includes a statement that describes the natural phenomenon to be explained, and the second part includes statements that represent general laws and antecedent conditions, which account for this phenomenon. The two parts of the explanation are closely related: in order for an explanation to be sound, statements describing phenomena (the first part of the explanation) must be logical consequences of general laws (the second part of the explanation) and the respective antecedent conditions.

Consider the following example: A person in a rowboat looking at the part of an oar that is under the water sees this part bent upwards. The phenomenon is explained by referring to some general laws, such as the law of refraction and the fact that air is optically less dense than water. According to the DN model, an explanation of this phenomenon is considered adequate when it refers to these laws and to some antecedent conditions such as the fact that part of the oar is immersed in water and another part is in air, and that the oar is a straight piece of wood. Thus, in an attempt to answer the question "Why did this phenomenon occur?" the answer includes "it occurred according to these general laws and in reference to these antecedent conditions." But what are these laws according to the DN model?

The concept of a law within the DN model. As clearly stated by Hempel, a DN explanation strongly depends on general laws. Hempel further stated that the absence of these laws renders an explanation invalid (Hempel, 2001). In the following section, a brief overview of Hempel's views on laws is presented. The overview starts with a science example that will help illustrate the role of laws in DN explanations.

Consider a piece of ice floating in a beaker of water at room temperature. As the piece of ice melts, one might wonder about the level of water in the beaker. Hempel briefly explained this phenomenon according to some general laws and principles. In fact, according to Archimedes' principle, a solid body in a container of liquid displaces a volume of that liquid that has the same weight as the body itself. And since melting ice does not alter the weight of the body, it turns into a mass of water that is of the same weight. Therefore, the level of water in the beaker does not change. The point to make here about laws is very important. The laws on which this account is based include Archimedes principle, the law of conservation of mass, and a law regarding the melting of ice at room temperature. None of these laws actually mentions this particular beaker of water or this particular piece of ice with which our explanation is concerned. Therefore, laws are not about this particular event only, but rather entail the general principle that under the same kind of circumstances, the same kind of phenomenon occurs. In addition, all laws that were used to account for this phenomenon can be also applied to the floating of a piece of stone in mercury or of a boat in water. Hempel makes an important distinction between the law that accounts for when any piece of ice floats in any water beaker and other laws that account for the phenomenon of any kind of solid in any kind of liquid container. Clearly, the former law seems weaker in that it deals only with the case of ice floating in water, whereas the other laws are more general.

Explanations vs. predictions within the DN model. In their covering law model, Hempel and Oppenheim (1948) were interested in the nature of scientific explanations and predictions, and how the two practices relate to each other. However, in their analysis the authors were mainly concerned with scientists, rather than students. Their views of predictions were mainly focused on the ability of a theory to predict an event prior to its occurrence. And that granted predictions more explanatory power: to be able to utilize a scientific theory or a general law in order to predict a certain phenomenon is surely a practice that experts do.

However, in science education the case is quite different. Students' predictions are not necessarily a result of a theory or a law, they are not always accurate, and more importantly they are not necessarily supported by evidence or a reason (e.g., Brewer, et al., 2000; Hogan & Maglienti, 2001). For the purpose of this study, scientific predictions are regarded as statements that posit the consequences of a phenomenon prior to its occurrence. It is important to note that although they play an integral part of the scientific practice, predictions are not automatically explanatory. A student's statement that the ball *will* fall to the ground when it is dropped is considered a prediction, and not an explanation. In addition, unlike descriptions, predictions are not necessarily or always based on observation. In many cases, they are based on students' prior knowledge and scientific background.

To illustrate how the difference between explanations and predictions is meaningful to science education, consider the first Performance Expectation (PS1) in middle school physical sciences in the NGSS (NGSS Lead States, 2013) as an example. PS1 is on matter and interactions:

> They [the students] will be able to provide molecular level accounts to explain states of matters and changes between states that chemical reactions involve regrouping of atoms to form new substances, and that atoms rearrange during chemical reactions (NGSS Lead States, 2013, p. 47).

An important dimension from the above statement is related to the expectation that students will *explain* changes between states. An example would be explaining *why* water changed to ice when it was kept in the freezer at 0°F for a certain period of time. A prediction, however, would entail predicting *what* would happen to water if it is kept in the freezer at 0°F for a certain period of time. In the example above, PS1 seems to aim at obtaining an explanation rather than a prediction: explanation of an event - that has already happened – by referring to some general laws and necessary conditions.

In high school physics science PS2, the NGSS (NGSS Lead States, 2013) seems to include both explanation and prediction without mentioning any differences between the two. PS2 is on motion and stability, and one of the questions that the students are expected to answer in this performance expectations reads “How can one explain and predict interactions between objects and within systems of objects?” (p. 75).

For Hempel, explanations are tools for understanding the world. Hence, understanding the world is a result of constructing explanations. But an integral question arises: what does understanding of the world entail? Hempel (1965) answered this question:

> Thus a DN Explanation answers the question '*Why* did the explanation-phenomenon occur?' by showing that the phenomenon resulted from certain particular circumstances, specified in C_1, C_2, …, C_k, in accordance with the laws L_1, L_2, …, L_k. By pointing this out, the argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred. (p. 337; italics in original)

The Inductive-Statistical (IS) model of explanation. While the DN model deals with explanations of deterministic structure, the IS model aims at probabilistic explanations. An important representation of IS explanations discussed by Hempel (1965, pp 385-386) is presented:

> Explanations of particular facts or events by means of statistical laws thus present themselves as arguments that are inductive or probabilistic in the sense that the explanation confers upon the explanandum a more or less high degree of inductive support or of logical (inductive) probability; they will therefore be called inductive-statistical explanations; or IS explanations.

Law-like statements in an IS explanation follow statistical laws of probabilistic nature. Hempel specified a condition of minimum degree of inductive support. He called it the *high probability requirement* (HPR), and required it to be high or closer to 1. Hempel did not specify a cut-off probability to HPR; however some philosophers (e.g. Weber et al., 2013) state that it must be higher than 0.5 (50% chance) for a statement to be a valid IS explanation. However, many philosophers have argued that HPR is neither necessary nor sufficient for valid statistical explanations (e.g. Salmon, 1998). The more important consideration when examining or constructing an IS explanation is to identify only the factors that are statistically relevant to the phenomenon-to-be-explained. So, even if there exists an outcome that is highly probable but unnecessary, then in some of these cases the improbable is more likely to occur: Even if a coin seems to be consistently biased for heads, it will still land tails-up.

Consider the following example provided by Salmon (1998) on this matter. Carbon 14 atoms decay in a statistically regular pattern providing a technique for radiocarbon dating. Other radioactive atoms decay with different statistical patterns. For example, the half-life of carbon 14 is 5715 years; the half-life of tritium (hydrogen 3) is 12.26 years; the half-life of uranium 238 is 4.46 billion years. One of the implications of these statistical regularities is that there exists a high probability that a given tritium atom, for example, will decay in a period of 5715 years – that is, there is 50% chance that a given carbon 14 atom will decay in the same period, and there is a smaller probability that a given uranium 238 atom will decay in that same period.

An interesting aspect of an IS explanation is that it includes phrases such as ‘it is practically certain that’ and ‘there is very high likelihood that’ instead of ‘a long series of repetitions’ or ‘approximately equal.’ The role of such phrases that are necessary in an IS explanation is to show the inductive, rather than the deductive nature linking statistical probability statements and empirical frequency statements involved in an explanation. Evidently, these phrases can be replaced by more quantitative phrases when applicable.

Many important explanatory accounts in science use statistical laws explicitly. In this regard, Hempel gives the following example: according to the Mendelian genetic principle, a random sample drawn from a population of pea plants, each of whose parent plants exhibits a cross of a pure-white flowered and a pure red-flowered strain, is highly probable to have red flowers (75%), and the rest would be white-flowered. Such a conclusion that may be used for explanatory, as well as predictive purposes, is in fact inductive-statistical. What it explains is the likelihood of obtaining red- and white-flowered plants in this sample. The high probability of obtaining red-flowered plants is due to (1) the relevant laws of genetics, some of which are statistical; and (2) the information of the type of the genetic make-up of the parent plants. The genetic principles that hold a universal (non-statistical) form include which color is dominant (red) over the recessive one (white), and others related to transmission by genes, etc. Additionally, the statistical generalization involved is the hypothesis that the four possible combinations of color genes – WW, WR, RW, RR – are statistically equiprobable. Accounts in terms of statistical laws have an integral role in science. They offer explanations that provide logical answers suggesting a different sense to the word ‘because’ (Hempel, 1965, p. 393).

Overview of the Covering Law model of explanation. According to Hempel and Oppenheim (1948), there seem to be at least two kinds of explanation: deductive and inductive explanations. The two kinds of explanations differ in two ways. First, even though both kinds demand the use of laws, deductive explanations require universal laws, which hold with no exceptions; whereas inductive explanations require statistical laws. In his book, Salmon (1998, p. 39) distinguished between universal and statistical laws. Universal laws are of the form “All F are G”, whereas statistical laws are generalizations of something that does not happen in every case, but rather in a specific percentage of cases.

The essay *Studies in the Logic of Explanation* by Hempel and Oppenheim (1948) is considered the seed from which almost all subsequent work on philosophical problems of scientific explanations stems. In the period of 1957 to 1958 a stream of works on scientific explanation began. These works were highly critical of the Hempel-Oppenheim covering law model. Criticisms of the covering law model came from Scriven (1958), Salmon (1984; 1998), and Kitcher (1989), among others.

Philosophical problems of the Covering Law model of scientific explanation. In what follows, the major problems of the DN and the IS models are discussed. The solutions to these problems shed light on the ways by which other models of explanation were developed.

The asymmetry problem: The flagpole example. Consider a flagpole of height H resting vertically on a flat surface on the ground. If the sun at a certain position is shining on the flagpole, then the latter casts a shadow of a certain length. When the height of the flagpole and the position of the sun are known, as well as the fact that light travels in a rectilinear path, one can deduce the length of the shadow. This deduction is in fact accepted by the DN explanation. Following the logic of the DN model, and given the facts about the position of the sun and the length of the shadow, one can also say that s/he can deduce the height of the flagpole. However, it is hard to accept that the height of the flagpole is explained by the length of its shadow. According to the DN model, knowing the facts and performing some mathematical calculations about the length of the shadow and the position of the sun can assist in deducing the correct height of the flagpole. According to Hempel's DN model, both derivations are equally explanatory.

An interesting point to mention is that both the causal and the unification models provide solutions to this asymmetry problem. While both Salmon's (and other philosophers) causal model and Kitcher's unification model are presented in more detail later in this chapter, it is worth discussing how both of these models offer solutions to this problem. Salmon (1989) argues that the reason why asymmetry is not accepted is because the flagpole *causes* a shadow and thus explains the length of the shadow; while the shadow does not *cause* the flagpole and hence cannot explain its height. As Carl Craver (2007) stated, causes explain effects; effects do not explain causes.

Hausman (1998) offered a solution to this symmetry problem of the DN model. He stated that only causal derivations are explanatory; derivations from effects are not. His criterion works as follows: Suppose that the angle of elevation of the sun is changed from 45° to 20°. Then the angle of elevation of the sun is independently changeable with respect to the other conditions specified in the derivation in the case that it does not affect them. This can be also applicable by changing the height of the flagpole. However, this does not hold for the flagpole's shadow. When the angle of elevation of the sun is changed, the flagpole's height does not change, but the length of its shadow does.

Finally, Philip Kitcher (1989) argued that the flagpole's height cannot be explained by the length of the flagpole's shadow. This is because, according to Kitcher, there is another derivation of the flagpole's height that represents an argument pattern with greater unifying power. Kitcher pointed out the fact that in the dark flagpoles have heights but not shadows. Therefore, explaining the height of the flagpole by the length of the shadow requires that one provides a valid explanation of the flagpole's height in the case it is dark. In so doing, there will be different explanations for the height of the flagpole when there is light and when there is not. Unlike Hempel, Kitcher believed that a scientist, presented with this problem, would consider the derivation of the length of the shadow from the height of the flagpole as explanatory, and the other derivation as non-explanatory.

Accidental generalizations. Hempel himself later realized a problem in his IS model: there seemed to be no distinction between what he called general laws and accidental generalizations. Consider the following example provided by Salmon (1989, p. 15) where he discusses this very problem:

(i) No gold sphere has a mass greater than 100,000 kg.

(ii) No enriched uranium sphere has a mass greater than 100,000 kg.

While the second statement appears to be that of a lawful fact, the first statement is a mere chance. Scientifically speaking, the critical mass of enriched uranium cannot be more than only a few kilograms; otherwise it would explode. On the other hand, the first statement is true only because no one has yet produced a sphere that heavy; there is enough gold in the world to make a 100,000 kg sphere that would not explode. The IS model does not offer any distinction between the two cases. More importantly, it does not state that the first statement is not a scientific explanation.

Irrelevant premises. Unlike Hempel's views, many philosophers of science believe that arguments with superfluous statements are not explanations. A famous example is provided by Salmon (1971, p. 34):

- (L) All males who take birth control pills regularly fail to get pregnant.
- (K) John Jones is a male who has been taking birth control pills regularly.
- (E) John Jones fails to get pregnant.

The problem with the above example is that the logic of it renders it an acceptable DN explanation regardless of the fundamental idea that males do not get pregnant.

In order to address the aforementioned and other problems with the covering law models, new philosophical models of explanation were developed. Causal models of explanations were among the first attempts at providing solutions to the Hempel's DN and IS models. In what follows, a brief overview of causality and explanation in philosophy of science is discussed.

Causality and Scientific Explanation

In this section, causal and causal mechanical explanations and their role in examining students' scientific explanation are presented in addition to the importance of manipulation in causality in science.

Causal explanations. Similar to Hempel's (1962) view of explanation, Salmon (1998) believed that scientific explanations are indeed answers to *why*-questions. However, Salmon added that not all 'why' questions are requests for scientific explanation. In particular, causal explanations are answers to 'why do/does?' rather than 'why should?' It is logical to affirm that usually causal explanations are derived through empirical investigations. There is usually a need to turn to causal explanations - through empirical investigation - in cases where deductive logic fails to answer our question about causality.

Consider the following example. After coming home from a day at the park, a child develops skin rash. On that day, different kinds of food were served in the park where different kinds of vegetation were present. In trying to find the cause of the child's rash, one might think that perhaps s/he had a large dish of strawberries, but also watermelon and apple, or played in an area of weeds. However, the single factor that was the cause of the rash is unknown, or whether or not the rash was a mere coincidence. Observations, however, can be made even to the point of conducting experiments in which the child eats strawberries but not watermelon or apple. The child can be given the food indoors away from plants. If the rash occurs every time the strawberries are given for example, but does not happen in the other circumstances, then a conclusion can be made that strawberries are the cause of the rash.

The above example clearly reminds any science teacher of controlled experiments. It also sheds lights on other factors, which are referred to as conditions of causality or as David Hume's features in causal situations: (1) the temporal precedence of the cause to the effect; (2) the spatiotemporal proximity of the cause to the effect; and (3) constant conjunction – the condition that every time the cause occurs, the effect follows. In the case of the child with the skin rash, eating a certain type of food preceded the skin rash; and it cannot happen the other way around. That is, the skin rash cannot precede its cause. Furthermore, the ‘space’ to look at for the cause of eating strawberries is the child’s body (rather than someone or something else). In addition, the skin rash should be noted within a reasonable amount of time after eating the strawberries (e.g., not a week after the child eats them). Finally, in order to be confident in stating that eating strawberries causes the child to develop skin rash, it has to happen every time the child eats strawberries.

David Hume’s three conditions of causality are useful while assessing students’ causal explanations. However, there are cases where they are not sufficient. Many philosophers argue that Hume missed an important condition for causal explanations. For instance, Salmon (1998) argued that Hume was “unable to find any ‘necessary connection’ relating causes to effects, or any ‘hidden power’ by which the cause ‘brings out’ the effect” (p. 85, quotations in original). Such a connection or a series of connections is what brings us to the Causal-Mechanical model of explanation.

The Causal-Mechanical model of explanation. Consider the following example from Salmon (1998): The ideal gas law does not emphasize any causal processes. In a gas container with a movable piston, when the gas is compressed by moving the piston – while keeping the temperature of the gas unchanged – the pressure increases. Such an increase in pressure can be explained causally drawing on a necessary connection of a lawful regularity. Traveling with the same average speed, the molecules collide with the walls of the container more frequently when the volume is less – since the walls are closer to each other. Note that the mathematical relationship between pressure and volume (at constant temperature) is not causal. However, the motion of the molecules obeying mechanical laws and colliding with the walls of the container are causal processes. Such causal processes that lead up to the event-to-be-explained are known as mechanisms.

Railton (1978) described explanations as statements that include causes and sequences of events that lead up to the event-to-be-explained. While Railton agreed with Hempel that explanations include references to law-like statements, he argued that they must also be supplemented by "an account of the mechanism(s) at work" (p. 748). Similarly, Salmon's (1984) work on causal mechanical explanation asserts that scientific explanations explain natural phenomena by showing how they fit in the causal structure of the world. Much of Salmon's work was on explicating what counts as causal processes and causal interactions. He defined a causal process as an entity that exhibits changes in its structure, and a causal interaction as an intersection among causal processes in which changes of the properties of these processes takes place.

While causal-mechanical explanations have come to refer mainly to the work of Salmon and Railton, other philosophers have developed different views on mechanisms. However, there seems to be a consensus that mechanisms are complex systems (Glennan, 2002). Glennan affirmed that "a mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations" (p. 344).

Let us now consider a causal mechanistic explanation. According to Glennan (2002), CM explanations are statements of mechanisms that include traditional accounts of explanations (e.g.

mechanisms that refer to general regularities or probabilistic laws) in addition to mechanistic systems. Consider for example a boy born with blue eyes. His father does not have blue eyes, but his mother does. Knowing that blue eyes is a trait of a recessive gene, one can say that the boy's mother must carry two copies of the recessive blue allele while the father carries one. Given the mechanisms of gamete formation, reproduction and the probabilistic laws of the genes responsible for eye color determination, there is a 50% probability that the boy will have blue eyes. This example is one of an IS explanation except that it does not meet Hempel's high probability requirement (a requirement that has been eliminated as necessary for an IS explanation to be valid). Therefore, instead of only referring the event-to-be-explained to a general or statistical law, a CM explanation explicates the causal processes and causal interactions that lead up to this event, in addition to subsuming it under a general or a probabilistic law and antecedent conditions.

Causality and manipulation. In his book *Making Things Happen: A Theory of Causal Explanation*, Woodward (2003) developed an account of causal explanation that is applicable to causal explanatory claims in various areas of science. Woodward drew on an important distinction between explanations and descriptions asserting that "views that take all forms of classification and description to be explanatory fail to satisfy this constraint" (p.5).

Biologists differentiate between description and classification on the one hand, and explanation and discovery on the other. In statistics, there exists a clear distinction between what is called descriptive statistics and inferential statistics, where the latter draws on causal relationships among variables under study. Hence, an adequate account of causal explanation should, Woodward (2003) argued, draw a clear distinction between descriptive information, and causal and explanatory information.

But what does such a contrast between descriptions and explanations entail? Robert Weinberg (1985) argued that in the past biology was regarded as a descriptive science. However, as it witnessed technological advances in instrumentation and experimental techniques, biology is now considered to provide explanations. Weinberg argued that the ability of biology to provide explanations lies in the fact that the information it delivers can be used for manipulation and control purposes. These advances have made it feasible to manipulate biological systems and observe results – an approach that was not possible in the past.

Woodward (2003) explicated the underlying idea of his account of causal explanation as follows:

> We have at least the beginnings of an explanation when we have identified factors or conditions such that manipulations or changes in those factors or conditions will produce changes in the outcome being explained. Descriptive knowledge, by contrast, is knowledge that, although it may provide a basis for prediction, classification or more or less unified representation or systemization, does not provide information potentially relevant to manipulation. (pp. 9-10)

Let us consider an example in order to illustrate Woodward's manipulative feature of causal explanations. The following is an example found in every high school physics textbook on laws of motion (see also Figure 2.1).

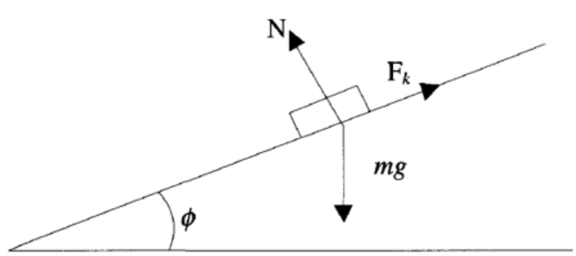


Figure 2.1 An example of a block in motion down an incline

Consider a block of mass **m** sliding down an inclined plane of angle **Φ** along the horizontal, with an acceleration **a**. How can the motion of the block be explained? As many physics teachers would state, the block is subject to three external forces: a gravitational force due to its weight (F_g), a normal force of support (N) perpendicular to the inclined plane, and a force of friction opposing the motion of the block (F_k). According to Newton's second law of motion, $F_k = \mu_k \bullet N$, where μ_k is the coefficient of kinetic friction. The gravitational force F_g is directed towards the center of mass of the earth and is equal to $F_g = mg \bullet \sin\Phi$. The normal force is $N = m \bullet g \cos \Phi$; hence the frictional force is $F_k = \mu_k \bullet m \bullet g \bullet \cos\Phi$. The net force on the block along the incline is the resultant of these two forces, hence: $F_{net} = m \bullet g \sin\Phi - \mu_k \bullet m \bullet g \bullet \cos\Phi$. And according to Newton's second law of motion, the acceleration of the block is given by: $a = g \bullet \sin\Phi - \mu_k \bullet g \bullet \cos\Phi$

In a science classroom and in a formative assessment of this kind, a student who applies Newton's second law of motion and correctly projects force vectors can obtain the last equation above. But Woodward (2003) asked how this so-called explanatory equation is in fact explanatory? His answer is again related to manipulation. Woodward argued that in order for the above account to provide explanation, there should be a set of relationships that are exploitable for manipulative purposes. More specifically, the above equations show us that if some variables are changed, the result that one is trying to explain would also change; i.e. the acceleration would change. For example, it is seen that if the angle of the inclined plane is increased, the acceleration of the block increases (and the block speeds down the incline). The final equation also shows us that if the mass of the block is greater (i.e., if the block is heavier), the acceleration will stay the same. This feature that Woodward introduces into causal models of explanation is an important criterion for explaining events (in this case, explaining the motion of the block) rather than explaining general laws. By and large, in this study the focus is on explaining events rather than explaining general laws. Seldom are cases of deriving or explaining natural laws encountered in K-12 science education. The focus is usually on students' explanations and understanding of events that obey general laws. For this purpose, Woodward's notion of causal explanations of particular events is adopted: they are invariant unless they are exploited for manipulation purposes.

Overview: Causation and scientific explanation. The topic of causation has always been of interest to philosophers, however there is little consensus on what causation really is. While to some philosophers causation is the incorporation of regularities and laws in cause-effect relationships, it is seen by others as manipulation and control. Another related reason for the little agreement is that different philosophers have different views about what counts as causal and what role necessary and sufficient conditions play in causal explanations.

A third and important reason is the connection between philosophical models of causation and developments in science. Recall Robert Weinberg's example on how such developments changed the nature of biology. Furthermore, in their introduction to the *Oxford Handbook of Causation*, Beebee, Hitchcock, and Menzies (2009) provided an interesting example in this regard:

> For example, Newton's celestial mechanics seems to posit instantaneous action at a distance, and quantum mechanics seems to tell us that the fundamental processes of our world are indeterministic. Both developments challenged existing assumptions about how causes could operate in our world. (p. 19)

This section discussed some of the main ideas of causality in scientific explanation that are relevant to science education. The three conditions of causality explicated by Salmon (1998) from the work of David Hume seem to be meaningful in science classrooms. It is important for science students to be aware of the temporal precedence of causes, spatiotemporal proximity of causes, and the necessity of the effect to happen when the cause occurs. Add to those conditions is the crucial feature of causality laid out by James Woodward: in order for an account to be considered causal, there should be a set of relationships that are exploitable for manipulative purposes.

The Unification Model of Explanation

Other models of explanation that also attempted to solve problems faced by the covering law model have also shown some problems posited by the causal model. For example, Philip Kitcher's unification model of explanation is another model that was developed in response to the covering law model. Kitcher (1989) argued that the causal account of explanation stems from a strong version of realism in which the world is viewed as having a certain structure that is independent of our effort to organize it. This section discusses a number of views of explanation as unification.

While not widely known, Carl Hempel's views of explanation actually aspired to unification. In his book, Hempel (1965) suggested:

> What scientific explanation, especially theoretical explanation, aims at is not [an] intuitive and highly subjective kind of understanding, but an objective kind of insight that is achieved by a systematic unification, by exhibiting the phenomena as manifestations of common underlying structures and processes that conform to specific, testable, basic principles (p. 83).

But what exactly does explanatory unification mean? As simply put by Feigl (1970): "The aim of scientific explanation throughout the ages has been *unification,* i.e., the comprehending of a maximum facts and regularities in terms of a minimum of theoretical concepts and assumptions" (p.12). In what follows, the work of Kitcher (1989) and Friedman (1974) on the unification account of explanation is discussed in more details.

The Unification Account of explanation. Unlike other accounts of explanation, the unification account posits an important definition of explanation: explanation is seen as an activity (an ordered pair that consists of a proposition and an act type) in which one answers questions to an audience drawing upon our beliefs using statements, laws, regularities, arguments, facts, theories, etc., that have been previously provided by scientists. According to this account, unification is accomplished by constructing arguments in which parts of our knowledge are derived from other parts. Kitcher stated that an argument is "a sequence of statements whose status (as a premise or as following for previous members in accordance with some specified rule) is clearly specified" (Kitcher, 1989, p. 431). In order for a statement to be considered explanatory – Kitcher referred to it as an argument pattern – it must be: (1) a sequence of schematic sentences; (2) a set of sets of filling instructions; and (3) a classification. A schematic sentence is a statement obtained by replacing some non-logical expressions in a sentence with dummy variables. Filling instructions are directions that replace the dummy variables. A classification depicts the inferential criteria of a set of schematic sentences.

An exemplar of unification is manifested by Newtonian laws of motion in such a way that one type of reasoning about certain mechanical principles can be used in the derivations of other phenomena. The unifying power, as Kitcher called it, of Newton's laws of motion is mainly in their ability to show how one set of scientific statements can be used over and over again in deriving other accepted scientific statements. Similar to Newton's laws of motion, Darwin's theory of evolution unifies various biological phenomena. Instead of merely providing a number of explanations of the existence of certain particular characteristics in some species, Darwin's theory shows how a pattern of scientific statements can be applied to a set of biological phenomena. More specifically, the theory demonstrates that by using certain scientific statements within a particular pattern, one can explain variations in similar species, variations of certain traits, characteristics of geographical distribution, and so on.

But one might wonder about the way in which derivations are obtained. Kitcher (1989) suggested an explanatory store. In this regard, he argued that science does not provide us with unrelated separate arguments and explanations, but rather it offers a reserve of explanatory statements – which constitutes the explanatory store:

> For a derivation to count as an *acceptable* ideal explanation of its conclusion in a context where the set of statements endorsed by the scientific community is K, that derivation must belong to the explanatory store over K, E(K). (p. 81)

So, instead of setting conditions on what every single explanation must satisfy, Kitcher suggests a *set* of explanations – the explanatory store – with the criterion of unification.

Friedman's views of unification. Friedman (1974, p.13) set clear properties that a theory of explanation should have. First, according to Friedman, a theory of explanation should be sufficiently general. Second, it should be objective in the sense that what counts as an explanation should not depend on the particular tastes and trends of a given historical period. Tastes and trends are non-rational factors, which have no place in an objective and rational sense of explanation that Friedman was after. Finally, it should connect explanation and understanding. The last property shows that Friedman did not want to do away with important psychological concepts such as *understanding*. Of course, Friedman still wanted to avoid explanations that would vary from individual to individual.

A brief overview of the unification account of explanation. Kitcher (1989) argued that major scientific breakthroughs have been accepted by communities due to their unifying power. He supported his view with examples such as Newton's and Darwin's theories. The difference between Kitcher and Friedman's views is related to the nature of explanations. While Friedman avoided explanations that would vary from one individual to another, Kitcher did not, and provided an example of the gun and the fusilier for this matter (1981, p. 510). Consider a fusilier asking Galileo about the reason why his gun shows maximum range (horizontal distance) when it is fixed on a flat surface and elevated at an angle of 45° with the horizontal. Galileo reformulates this question in his mind into a basic question of ideal projectile motion that when having a fixed velocity and subjected to only gravitational acceleration, shows a maximum range at angle 45° neglecting the effects of air resistance. Now Galileo adapts his arguments and explains to the fusilier the familiar terms by eliminating computational steps. The result is that both Galileo and the fusilier are satisfied. Kitcher argues that what Galileo does is that he selects an explanation from the explanatory store and reformulates it based on the context.

The Pragmatic Model of Scientific Explanation

When a teacher asks a fourth grader "why does this piece of metal expand?", a typical answer might contain something like "this metal was heated; and all metals expand when heated." This answer is considered acceptable when provided by a fourth-grade student. However, based on prior knowledge and student level, the same answer would not be as acceptable (or even 'correct') if it is provided by an undergraduate chemistry student for example. Thus, what counts as an adequate scientific explanation changes by students' level, prior knowledge and other factors determined by the teacher. This case simplifies the idea of the pragmatic model of scientific explanation. For a more accurate and philosophical elaboration of the pragmatic model of scientific explanation, there is a need to delve deeper into the ideas explicated by developers and supporters of this model of explanation.

Hempel's pragmatic views of explanation. Carl Hempel broke new ground in his work on scientific explanation not only because he was the first to develop a model on scientific explanation (the covering law), but also – as discussed earlier in this chapter– because he brought up ideas on explanation that were later developed in the philosophy of science (such as the unification model of scientific explanation). In fact, Hempel (1965) expressed views on the pragmatic nature of scientific explanation. He wrote:

> Very broadly speaking, to explain something to a person is to make it plain and intelligible to him, to make him understand it. Thus construed, the word 'explanation' and its cognates are pragmatic terms: their use requires reference to the persons involved in the process of explaining. In a pragmatic context one might say, for example, that a given account A explains fact X to person P_1. It is noteworthy then that the same account may well not constitute an explanation of X for another person P_2, who might not even regard X as requiring an explanation, or who might find the account A unintelligible, or unilluminating, or irrelevant to what puzzles him about X. (p. 425)

At a first glance, Hempel (1965) is seen to acknowledge such a pragmatic nature of explanation. However, in his work on developing the covering law model, he saw his own task as "constructing a non-pragmatic concept of scientific explanation – a concept which is abstracted, as it were, from the pragmatic one, and which does not require relativization with respect to questioning individuals" (p. 426).

In this regard, Peter Achinstein (1984) interpreted Hempel's view of the pragmatic character of explanations as follows:

(1) Account A explains fact X to person P

Statement (1) emphasizes a reference to a certain person who is explaining or receiving an explanation – and this form is in fact pragmatic.

However, Hempel also seems to be saying, as Achinstein interpreted it, that there is another form of explanation such that:

(2) Account A explains fact X.

Statement (2) does not include a reference. Such form is a non-pragmatic type of explanation – just as Hempel's covering law is.

Therefore, it seems that Hempel's characterization of pragmatic explanations included explanation-sentences that are pragmatic if they contain terms of a certain explainer or audience. Hempel also made a distinction between explicit and implicit pragmatic explanations – the explainer/audience is not necessarily explicitly stated so that the sentence is pragmatic. Nonetheless, Hempel acknowledged yet another type of explanation – in his covering law model – that is of non-pragmatic nature.

Van Fraassen's pragmatic views of explanation. While there appears to be a philosophical view that there are two types of explanations (pragmatic and non-pragmatic), some philosophers necessitated a reference/audience for an adequate scientific explanation. In his book on the *Pragmatic Theory of Explanation*, Bas van Fraassen (1980) directly opposed Hempel's views on pragmatism. Van Fraassen wrote:

> The description of some account as an explanation of a given fact or event, is incomplete. It can only be an explanation with respect to a certain relevance relation and a certain contrast-class. These are contextual factors, in that they are determined neither by the

totality of accepted scientific theories, nor by the event or fact for which an explanation is requested. (p.130)

Van Fraassen clearly considered the statement “Account A explains fact X” to be incomplete. Consider the following example of a scientific explanation that applies to van Fraassen’s pragmatic view: In elementary earth and space science (ESS) in NGSS, ESS1 on Earth’s place in the universe includes an assessment of the fact that “star pattern is limited to stars being seen at night and not during the day” p. 14. A first-grade teacher shows a picture of stars in the sky during night, and asks *Why are there stars in the sky at night?* One can interpret this question in different ways. For example, the teacher might be asking:

- Why are there stars (rather than something else) in the sky at night?
- Why are there stars in the sky (rather than somewhere else) at night?
- Why are there stars in the sky at night (rather than during the day)?

Sometimes these distinctions are implicit and understood from the context of learning. However, whether implicit or explicit they are important for accuracy, and they target a request for a complete scientific explanation.

Achinstein’s pragmatic views of explanation. Peter Achinstein’s (1984) pragmatic model of explanation is concerned mainly with explanation-sentences of the form:

S explains q by uttering u, where q is the indirect form of a question Q.

In this regard, Achinstein (1984) noted that “such sentences are true if S utters u with the intention of rendering q understandable by producing the knowledge that u expresses a correct answer to the question Q” (p. 282). Three important implications can be drawn from Achinstein’s pragmatic account. First, the above explanation-sentence necessitates a reference to an explainer. Second, the fact that the intention is understanding emphasizes the pragmatic –

mainly the subjective – nature of scientific explanations. Finally, Achinstein was not only concerned with the context in which the explanation takes place, but also with the pragmatic nature of the content of explanation itself.

In an aim to elaborate the third and important implication of his pragmatic account, Achinstein provided an excellent example of the added value of the pragmatic notion to scientific explanation. In what follows, a summary of his example of the atomic model is presented, and a discussion of how his model offers new insight on scientific explanation is followed (for the interested reader, the full example can be found in Achinstein, 1984, pp. 284-288). Before discussing the example, let us first discuss Achinstein's views on what scientific explanation is.

Achinstein (1984) stated that "an explanation of q can be construed as an ordered pair whose first member is a proposition or set of propositions that constitute an answer to Q, and whose second member is a type of explaining act, viz., explaining q" (p. 282). Achinstein's pragmatic definition of explanation differs from other definitions in traditional accounts in that an explanation does not need to be an answer to a "why?" question. According to Achinstein, an explanation could be of an answer to a "what?" and a "how?" event. He thus replaced *why*-questions with a more meaningful criterion to explanations: content-questions. Achinstein further believed that there are several evaluations of explanations depending on their aim. An aim might be purely universal (such as the achievement of truth), or it might be contextual. However, the aim with which Achinstein's pragmatic model is concerned is one that an explainer has when s/he provides an act of explaining q that makes it understandable to a reference audience by producing the appropriate/correct answer to a question Q. Such a pragmatic evaluation takes into account both contextual, as well as universal aims.

In 1901 Geiger and Marsden published their experiments showing that when alpha particles are directed towards a thin metal foil, most of them go through the foil while some get scattered bouncing back. In order to explain these unexpected results, Rutherford proposed a new theory of the structure of the atom. In his theory, Rutherford assumed that the atom contains a positive charge distributed unevenly but concentrated in the nucleus. He also assumed that the moving electrons surround the positively charged nucleus. Finally, he assumed that scattering was due to the encounter between the alpha particles and a foil atom. Since most alpha particles were able to go through the foil without being scattered, the foil atoms were mostly empty of matter. From these assumptions and the principles of classical mechanics including the conservation of energy and momentum, Rutherford derived a formula that gives the number of alpha particles falling on a certain area and deflected at a certain angle in terms of other quantities.

Achinstein (1984) asked if Rutherford's explanation of the scattering of alpha particles is a good explanation? He then answered his question arguing that if one examines this explanation in terms of non-pragmatic criteria, then they will get mixed reviews. Rutherford's explanation derives the angles in a precise way from law-like quantitative assumptions, and offers a cause of the scattering. Nonetheless, it turned out later to be only an approximation to what really happens in the foil atoms. But is it still a *good* explanation? According to Achinstein (1984), if one uses only non-pragmatic criteria in our examination of the explanation, it will be hard to say why Rutherford's explanation is better than other explanations. Consider for example Geiger and Marsden (G-M) quantitative hypothesis that they developed without using Rutherford's theory. Without delving into the mathematical aspect of the G-M hypothesis, it also explains the scattering of alpha particles in a precise manner using law-like quantitative assumptions. In addition, the G-M hypothesis is unifying since it allows the derivation of other results obtained in other similar experiments. It is also causal in that it mentions the causes of scattering. Yet, why is it regarded as inferior to Rutherford's explanation?

Achinstein (1984) argued that non-pragmatic criteria such as derivability from laws, unification, and causation are not *by themselves* sufficient to tell us why one explanation is better than another. In fact, Rutherford's explanation is better than the G-M explanation not because it answers a causal question in a unifying law-like manner, but because "it does so at the subatomic level of matter in a way that physicists at the time were interested in understanding the scattering" (p. 286). Achinstein further believed that the reason the G-M explanation is not as good is that it did not appeal to what physicists were looking for at the time.

The above example highlights the importance of pragmatism from a content perspective. However, the pragmatic model that science educators might be more interested in is the one that is more contextual in nature. Nonetheless, Achinstein's (1984) model adds this insightful criterion to scientific explanation: the importance of the kind of laws, causation and unification the explainer is looking at when providing a complete explanation. In fact, what makes a pragmatic explanation better than a non-pragmatic one, according to Achinstein, is the idea that objective non-pragmatic explanations will always be faced with counter-explanations. Therefore, in order to construct meaningful explanations, students will still have to use laws, causal factors and unification. But unless they say something *more specific* about these laws and causal factors, or what needs to be unified, their explanations will not be complete.

Overview of the pragmatic account of explanation. This section presented views on the pragmatic nature of explanation from Hempel, van Fraassen and Achinstein. While Hempel acknowledged contextual explanations of pragmatic nature, he gave more importance to non-pragmatic explanations. On the other hand, van Fraassen and Achinstein emphasized the need of a reference class/audience for a complete scientific explanation.

Table 2.1 summarizes the various philosophical models of explanations discussed so far. It includes general definitions of each form of explanations, examples, and other related issues. One very important criterion of this table is that each form of scientific explanation takes place in a contextual/pragmatic medium. This determines the nature, structure and form of each explanation. The emphasis on contextualization in explanation is what drives a meaningful K-12 schema for studying scientific explanation.

A Pragmatic Approach to Studying Scientific Explanation

The work of Weber et al. (2013) on explanation in the social and biomedical sciences has resulted in the development of an alternative approach to studying scientific explanation. Their approach is pragmatic in a sense that it uses one or more of the traditional philosophical models as tools to studying scientific explanations.[1] Furthermore, the approach focuses on examining how and why scientists ask explanation-seeking questions. While this approach is more philosophical, some of its elements are applicable to the science classroom.

Weber et al. (2013) have developed a toolbox of questions and answers formats for examining scientific explanation. For the purpose of this study, only the parts of the toolbox relevant to science education are included. For instance, in their toolbox Weber et al. focus on explanations of the reasons behind certain actions/behaviors of scientists (related to ethics, knowledge, and ultimate truth). Such aspects of the toolbox are out of the scope of this study. Hence, in the following section, the relevant elements of their toolbox for studying scientific

1 Weber et al.'s (2013) pragmatic approach to explanations should not be confused with the pragmatic account of explanations (e.g., van Fraassen, 1980). One of the most important criterion of the approach is the use of the philosophical accounts as tools to examine explanations, while the pragmatic models of explanation focuses on the contextual aspects with which scientific explanations are constructed.

Table 2.1

Definitions, Conditions, and other Related Issues of Philosophical Models of Scientific Explanation[2]

Philosophical Model	Definition	Example	Conditions or Criteria	Problems & Solutions
Deductive-nomological (Hempel & Oppenheim)	*DN explanation* consists of two parts: statements that describes the natural phenomenon to be explained, and statements that represent general laws that account for the phenomenon. The first part must be logical consequence of the second part.	General example: Why did this phenomenon occur? It occurred according to these general laws and in reference to these necessary conditions.	*General laws* imply the general principle that under the same kind of circumstances, the same kind of phenomenon occurs. They must be of essentially generalized form so as to be able to serve an explanatory role.	Asymmetry problem: Use DN models only in cases where causes explain effects (effects do not explain causes).
Inductive-Statistical (Hempel & Oppenheim)	*IS explanation* confers upon the explanandum a high degree of inductive support of logical probability. IS explanations are evaluated by expectability that comes in degrees/probabilities.	Flipping of a coin	*Higher probability requirement (HPR)* requires a probability higher than .5 *Requirement of maximal specificity (RMS)* an IS explanation should include all relevant information (and avoid irrelevant information).	-Accidental generalizations vs. general laws -Irrelevant premises
Causal (Salmon)	*A causal explanation* asks why do/does rather than why should. We usually turn to causal explanations through empirical investigation.	Controlled experiments	*David Hume's conditions of causality*: (1) temporal precedence of the cause to the effect, (2) spatiotemporal proximity of the cause to the effect, and (3) the condition that every time the cause occurs the effect follows.	-Quantum mechanics -Evolutionary biology

(Table continues)

[2] All of these philosophical models of explanation take place in a pragmatic context. More specifically, the structure and content of explanation is subject to change based on students' level, their prior knowledge, and other subjective/qualitative criteria that depend on teacher judgment, context of curricular development, etc. It is important to keep in mind that an explanation of the question 'Why did a phenomenon A occur?' varies in content and structure between a 4th grader and a high school student, even though in both cases it might be a DN explanation, for example.

Table 2.1 (continued)

Philosophical Model	Definition	Example	Conditions or Criteria	Problems & Solutions
Causal Mechanical (Salmon)	Instead of only referring the event-to-be-explained to a general or statistical law, a CM explanation explicates the causal processes and causal interactions that lead up to this event, in addition to subsuming it under a general or a probabilistic law and antecedent conditions.	The ideal gas law example	*David Hume's conditions of causality*: (1) temporal precedence of the cause to the effect, (2) spatiotemporal proximity of the cause to the effect, and (3) the condition that every time the cause occurs the effect follows. PLUS: (1) It must explicate all necessary causal connections (in the form of processes and interactions) that lead up to the event-to-be-explained.	Quantum mechanics Evolutionary biology
Causality, Manipulation and Control (Woodward)	In order to provide a valid explanation, there should be a set of relationships that are exploitable for manipulative purposes	A case of a block down an inclined plane	*Explanation vs. Description* The ability to move from providing a description to providing an explanation lies in the fact that the information it delivers can be used for manipulation and control purposes.	Not all cases are exploitable (e.g., geology and astronomy).
Unification account of explanation (Kitcher)	*Unification* takes place by exhibiting the phenomena as manifestations of common underlying structures and processes that conform to specific, testable, basic principles.	Newtonian Mechanics	Explain the maximum number of facts in terms of the minimum possible number of theoretical concepts, general laws and assumptions. *The explanatory store* Science offers a reserve of explanatory statements that constitutes the explanatory store.	Kitcher's account seems to promote information compression—deriving as much from as few patterns of inference as possible.
Unification account of explanation (Friedman)	According to this account, a theory of explanation should be general, objective, and should lead to understanding.	The gun and the fusilier	The link between explanation and understanding emphasizes the importance of the psychological concepts that are closely associated with explanation.	Friedman avoids explanations that would vary from one individual to another, Kitcher does not.

explanations in the science classroom are summarized. In addition, examples, mainly from the NGSS Lead States (2013),that are applicable to elements of the toolbox are provided. The pragmatic approach and the toolbox to studying scientific explanations along with the traditional philosophical models of explanations were the guiding principles of the development the NOSE framework.

Toolbox for Studying Scientific Explanations

As mentioned earlier in this chapter, there is a general agreement regarding the idea that an explanation consists of two parts: the first part includes statements about *what* or *how* something is, and the second part includes statements about *why* something is. Using the pragmatic approach, Weber et al. (2013) developed what a toolbox for examining each part of a scientific explanation. The toolbox includes a distinction between explanations of plain facts and explanations of regularities/laws. The present study is concerned with the former type of explanations. Rarely are explanations of laws and regularities required in K-12 context. Instead of asking “Why does light obey the law of refraction when passing from one medium to another?”, teachers usually ask “Why does light bend when passing from one medium to another?” In the following section, Weber et al.’s toolbox for questions and answers format is summarized.

Questions about plain facts. The simplest form of questions about plain facts is “Why is it the case that A?” where A is a statement about a particular fact. This type of question is commonly asked in the science classroom. A quick look at the NGSS document reveals many examples of this type: “Why do some objects keep moving?”; “Why do objects fall to the ground?”, among others. Furthermore, questions about plain facts have been examined, as has been discussed earlier, by Hempel, Kitcher, and Salmon. Hempel (1965) believed that all explanation-seeking questions about particular facts are of this format. He wrote:

> [A]nd in that case the explanatory problem can again be expressed in the form 'Why is it the case the *p*?', where the place of '*p*' is occupied by an empirical statement specifying the explanandum. Questions of this type will be called *explanation-seeking why-questions.* (p. 334)

Making a similar assumption as Hempel, Philip Kitcher (1981) wrote:

> To determine the condition under which an argument whose conclusion is S can be used to answer the question "Why is it the case that S?" (p.510).

Salmon made the same assumption, but implicitly. In fact, the examples he provided fit this format and he regularly used the term "the fact-to-be explained" when discussing the explanandum (see, Salmon 1984, p. 13, and pp. 15-19).

Answer formats to questions about plain facts. Weber et al. (2013) suggested five answer formats to *why*-questions about plain facts: Causal deductive nomological (CDN) formats, causal inductive statistical (CIS) formats, causal default rule (CDR) formats, positive causal relevance (PCR) formats, and positive and negative causal factors (PNCR) formats.

The causal deductive-nomological (CDN) format. A CDN answer format (1) includes an explicit relevant population; (2) must be purely inferential, meaning it makes no assumptions about evidence for causal claims; and (3) highlights the practical usefulness of causal knowledge. Consider the following example: If we assert that smoking leads to (is a positive causal factor, or PCF, of) lung cancer in a certain population, this supports the assertion that if every person in this population was forced to smoke then there would be more people with lung cancer than if every person in this population was forbidden to smoke. Conversely for the assertion that smoking does not lead to (is a negative causal factor, or NCF, of) cancer. Finally, if we claim that smoking is causally unrelated with the occurrence, or otherwise, of cancer, then we can claim that in the two hypothetical populations above the probability of having cancer is equal.

The causal inductive statistical (CIS) format. As discussed above, Hempel developed two models: the DN model and the IS model. Similar to adopting a DN model and adding a causality requirement to it (thus developing the CDN format above), the CIS model is adding a causality requirement to the IS model.

The causal default rule (CDR) format. This format is a variation of the CDN and CIS formats in which there is no requirement for deduction or induction with an exact probability value (as is the case with IS explanations). Instead, it is based on inductive inference of what Weber et al. (2013) called default rules. Default rules (such as ‘Birds usually fly’) are different from universal laws that are used in DN explanations. Default rules allow exceptions (e.g., ‘Penguins do not fly’). In addition, they are different from probabilistic statements characteristic of IS explanations in that they do not require specific probabilities for exceptions and cases of normality (or non-exceptions). In most cases, non-exceptions can include occurrences fairly close to probability = 1. If one can specify the relative frequency, then the explanation becomes an IS explanation. However, information needed to attain relative frequencies is not always available. Thus, CDR explanations are considered as a third possible form of covering law explanations.

The positive causal relevance (PCR) format. Adapted from the work of Nancy Cartwright (1983) on causality, the PCR format necessitates that a valid explanation must contain at least one PCF; otherwise the explanation is inadequate. Cartwright’s poison oak example that has been discussed earlier in this chapter provides a good elaboration of PCR format:

> Spraying poison oak trees with defoliant X increases the probability of their death. The poison oak tree in my garden died because I have sprayed it with defoliant of brand X.

The above explanation represents the simplest form of a PCR format: one PCF and no NCF. While Cartwright claimed that PCR is the only format for an adequate scientific explanation, it is considered as only one tool in this toolbox.

The positive and negative causal factors (PNC) format. As has been discussed earlier in this chapter, the PNC was adapted from Humphreys' (1989) idea that both PCF and NCF are conditions for an adequate explanation. Humphreys' plague example is a good example of the PNC Format:

> The bubonic plague bacillus (Yersinia pestis) will, if left to develop unchecked in a human, produce death in between 50 and 90% of cases. It is treatable with antibiotics such as tetracycline, which reduces the chance of mortality to between 5 and 10%. (Humphreys, 1989, p. 100)

Explaining Albert's death who contracted the bubonic plague, Humphreys writes:

> An appropriate response at the elementary level would be "Albert's death occurred because of his infection with the plague bacillus, despite the administration of tetracycline to him." (Humphreys, 1989, p. 100)

According to Humphreys, an explanation that leaves out the tetracycline (which is the negative causal factor in this case) is inadequate since it made Albert's death less probable.

Transforming the above explanation into the PNC format, we get:

Albert was infected with the plague bacillus.

Albert took tetracycline.

Albert, who is a human, is dead.

Contrastive questions. As presented earlier in this chapter when discussing his pragmatic model of explanation, van Fraassen (1980) cautioned against the use of questions about plain facts. In fact, van Fraassen argued that all *why*-questions must be contrastive of the form "Why A rather than B?". Van Fraassen did not deny the fact that people (including scientists), ask non-contrastive why-questions. However, he argued that such questions are inaccurate formulations of contrastive questions because the real problem that needs to be answered lies in the contrast class. Although pointing out the difficulty in proving such a claim, Weber et al.'s (2013) toolbox supported it in two aspects, writing:

(1) Many *why*-questions that scientists ask are contrastive in nature.

(2) If a scientist asks a non-contrastive question, it is *sometimes* the case that this question does not adequately represent the cognitive problem of the scientist: the real problem h/she wants to tackle is contrastive. (p. 41)

Note that for this toolbox there is an important shift from van Fraassen's *all* to Weber et al.'s *sometimes* in claim (2) above. Because both contrastive and non-contrastive *why*-questions appear in scientific practice, Weber et al. included both types in their toolbox. The first type is simply a question about plain facts.

In the NGSS document, we find questions of both types. Nonetheless, explicit contrastive questions appear less infrequently. I fact, the vast majority of *why*-questions in the NGSS are non-contrastive, or as van Fraassen would argue, implicitly contrastive. For example, on page 79 of the NGSS State Lead (2013), the question "How and why is the Earth constantly changing?"

can be rephrased as “How and why is the Earth constantly changing rather than remaining the same?”. Another example on page 4 reads “Where do animals live and why do they live there?” can be expressed as “Where do animals live and why do they live there rather than somewhere else?”.

While one can transform almost any question of the form “Why is it the case that A?” into “Why is it the case that A rather than B?”, there might be multiple alternatives. For instance, recall the above question “Where do animals live and why do they live there?”. This question can be expressed as “Where do animals live and why do they live there rather than somewhere else?”. But it can also be expressed as “Where do animals rather than other species live and why do they live there?”. Hence, contrastive questions target the scientific concept that is being questioned more than what questions about plain facts do.

Does that mean that all questions about plain facts should be transformed into explicit contrastive questions? For the most part, the answer is yes. However, in many cases the teaching context determines what the alternative part in a particular question should be. This takes us back to the importance of the contextual/pragmatic nature of scientific explanation – a nature that should be taken into account in every form of scientific explanation.

Answer formats to contrastive questions. According to Weber et al.’s toolbox, answers to contrastive questions should inform us, in some way or another, how the world should have been different if the alternative were to occur. Consider the following question: “Why does a plastic ruler attract small pieces of paper?” can be expressed as “Why does a plastic ruler rather than a wooden rod attract small pieces of paper?”. Another related question can be “Why does a plastic ruler attract small pieces of paper rather than small pieces of leather?”. Recall that the simplest form of the contrastive question “Why A rather than B” indicates that A has occurred while B has not, but might be possible. In what follows, a discussion of the two answer-formats for this type of questions in this toolbox are presented.

Reality to alternative format[3]. Answering questions of the form “Why A rather than B?” requires first describing A (what really happened), explaining A (why it happened), and then presenting B (an alternative scenario that would have produced a different result if it had occurred).

Recall the question “Why does a plastic ruler rather than a wooden rod attract small pieces of paper?”. When the plastic ruler is rubbed on hair, it acquires a net negative electric charge. This is because upon rubbing (a form of charging/electrification), plastic tends to gain electrons (hence becoming negatively charged). A piece of paper is neutral, i.e. it is made up of electrically neutral molecules. However, this does not mean that it lacks electric charge; molecules have positively charged nuclei and a negatively charged electron cloud. So, when a negatively charged plastic ruler approaches a piece of paper, the electrons that are on the plastic ruler repel those that are on the tiny pieces of paper (like charges repel) causing a redistribution of charges. The positive charges on the pieces of paper are now closer to the negative charges on the plastic ruler, thus causing an attraction (unlike charges attract). On the other hand, when a wooden rod is rubbed on hair, the rod might gain very few electrons but will stay mostly neutral (due to the nature of wood). Hence, a neutral wooden rod that is previously rubbed on hair remains neutral. When it is held close to the neutral pieces of paper, nothing will happen.

As can be seen from this example, in reality to alternative explanation formats, we describe what really happened and explain why it happened – then we describe what would have

[3] Weber et al.’s (2013) toolbox includes another answer format for contrastive questions – an alternative to reality format – that is concerned with the reasons behind someone choosing a scientifically incorrect answer to a problem. While important, this format is not relevant to formulating a scientific explanation. It tackles issues such as circumstances that lead to a wrong answer, inadequate skills, etc.

happened in an alternative scenario, and describe why it could not have happened. In a science education context, such an example can be hands-on (where students experiment the reality and alternative scenarios). The gist of such explanation formats is offering description of each scenario and explaining why each scenario really happened (or could not have happened).

Real contrasts format. The second type of answer formats to contrastive questions are real contrasts formats. In science, we often encounter problems that have two or more alternatives. Consider the example of the plastic ruler and the pieces of paper in which one alternative (having a wooden rod rather than a plastic ruler) was discussed. In real contrasts explanations, one might consider other alternatives (such as an answer to the question "Why does a plastic ruler attract small pieces of paper rather than small pieces of leather?", or "Why does a plastic ruler rather than a glass rod attract small pieces of paper?". Hence, real contrasts formats have to include descriptive and comparative parts, where the latter lists the differences between the descriptive parts. It must also include one or more alternative scenarios that inform us how the world would be different if the alternative scenarios were to take place.

Resemblance questions. Weber et al. (2013) argued that resemblance questions have not been of interest to philosophers examining scientific explanation even though they occur in scientific practice. An example of a resemblance question is: "How are parents and their children similar and different?" (NGSS State Lead, p. 10), or "How are plants, animals, and environments of the past similar or different from current plants, animals, and environments?" (p. 22). Evidently, resemblance questions focus on similarities between events as opposed to only differences (contrastive questions) or on a single event (questions about plain facts). In the NGSS, we find that all resemblance questions are also contrastive; i.e., they seem to focus on similarities and differences between events.

Answer formats to resemblance questions. Stemming from the fact that resemblance questions target similarities between events, explanations that answer this form of question focus on unifying these events. Unification in this context refers to what two (or more) events have in common. Weber et al. (2013) considered two answer formats for resemblance questions: Top-town unification format and bottom-up unification format. Before summarizing each of these answer formats, it is important to discuss how these formats (or so-called tools) are derived from (or inspired by) philosophical models of explanation.

Earlier in this chapter, Kitcher's unification model of explanation was summarized. More than a decade after Kitcher, Robert Skipper (1999) provided an alternative to Kitcher's unification model of explanation. This alternative is relevant to explanations as answers to resemblance questions. Skipper (1999) argued:

> I have provided the foundations of an alternative to Kitcher's way of understanding explanatory unification. Kitcher claims that unification is the reduction of types of facts scientists must accept in expressing their worldview, and it proceeds through derivation of large numbers of statements about scientific phenomena from economies of argument schemata. I suggest that it is very much worth exploring whether unification can be conceived as the reduction of types of mechanism scientists must accept as targets of their theories and explanations, and whether it proceeds through the delineation of pervasive causal mechanism via mechanism schemata (pp. 207-208).

Weber et al. (2013) assert that while top-down unification format is vital to scientific explanation, bottom-up unification format is almost always non-scientific, but rather related to societal behaviors. In the following section a brief discussion of the bottom-up unification account is presented.

Top-down unification format. In the top-down unification format, events to be unified belong to the same law (or set of laws) of nature. Explanations of this format incorporate different events under one law (or set of laws). Accordingly, when formulating a top-down unification explanation, statements are constructed in a way that show that the events are to be expected from the same law (or set of laws) that is included. Weber et al. (2013) provide the following example:

Why do Peter and Mary both have blood group A?

This question can be answered as follows:

Unifying Law:

All humans who belong to category I_AI_A x I_AI_O have blood group A.

Application 1

All humans who belong to category I_AI_A x I_AI_O have blood group A.

Mary is a human and belongs to category I_AI_A x I_AI_O.

Mary has blood group A.

Application 2

All humans who belong to category I_AI_A x I_AI_O have blood group A.

Peter is a human and belongs to category I_AI_A x I_AI_O.

Peter has blood group A.

This example is one of a top-down explanation. It shows that both events are to be expected based on one law. Such cases (two events, one law) form the simplest top-down unification formats. In science, more than two events are encountered and more than one law can be used. As long as the same laws for all events are used, then a top-down unification explanation format can be applied. For instance, consider the set of two laws stated in the NGSS

(2013) in high school disciplinary core ideas: Newton's law of universal gravitation and Coulomb's law of electrostatics. This set of laws is used, as the NGSS document states "to describe and predict the effects of gravitational and electrostatic forces between distant objects" (p. 84). Several events are derived from both of these laws. For example, the force that holds us to the surface of the earth and prevents us from floating into space; the force that causes the earth to orbit around the sun in an elliptical motion; the net charge of the universe can neither decrease or increase; and finally we can calculate the force between any two charged objects. All these events, and more, can be explained by the Newton's law of gravity and Coulomb's law of electrostatics. Kitcher's (1989) view of unification stems from the idea that unification is attained by formulating deductive statements that belong to acceptable patterns. Kitcher's ideal scenario is to construct as many explanations as we can (for as many natural phenomena as possible) using the least number of patterns. A noteworthy criterion that Weber et al. (2013) added is the requirement that top-down unification should always be based on causal laws.[4]

Bottom-up unification format. Bottom-Up unification answer format addresses explanations of how different events share identical causal factors. The common factors that provide the unification of several events are not enough for causing these events. Hence, one cannot consider an explanation of these events to be a top-down, but rather a bottom-up unification. This is because the mechanisms that lead to the events include identical causal factors. What makes this answer format nonscientific is that the common factors in this explanation format do not need to be derived/deduced from laws. Table 2.2 summarizes the

[4] The reader can refer to Weber and Van Bouwel (2009) in which the authors argue that it is nearly impossible to find an acceptable non-causal explanation – hence adding the condition of causality to almost all explanation formats.

elements of the toolbox discussed in this section. It includes the three types of questions, their corresponding answer formats and the philosophical models that guided them.

Advantages of the Scientific Explanation Toolbox

Weber et al.'s (2013) toolbox is useful in the field of science education in many aspects:

- *The toolbox is guided by philosophical models of explanation*. It covers almost all philosophical models of scientific explanation, thus offering a clear set of guidelines and modalities that are specific to scientific explanation.
- *The toolbox focuses on the content and the structure of scientific explanations*. It portrays explanation-sentences in the form of answer formats that focus on the content and structure of scientific explanations based on philosophical perspectives. For example, all relevant answer formats to questions about plain facts include causality in one form or another. That is to say that according to this toolbox, a valid explanation to a *why*-question about plain fact must be causal. We find explanations combining deduction with causality (CDN), induction with causality (CIS), positive and/or negative causal relevance (PCR and PNC), and so on.
- *The toolbox uses philosophical models as tools to studying scientific explanation***.** Answer formats of resemblance questions are abundant in biology (recall the blood group example). This toolbox shows that some explanations are neither deductive/inductive nor causal. Instead, they rely on unifying laws of nature. This aspect of the toolbox emphasizes a meaningful use of the philosophical models as tools to studying scientific explanation.

Disadvantages of the Scientific Explanation Toolbox

While the toolbox offers a unique approach to studying scientists' explanation in a meaningful way, it lacks some elements that are important in K-12 science classrooms:

- *The toolbox does not consider the explainer's prior knowledge, level or context of learning.* Evidently, the toolbox aims at analyzing scientists' explanations. Therefore, their decision could be due to the fact that their audience is known and is somewhat invariant. However, an audience reference is of crucial importance in a schema for analyzing K-12 students' scientific explanations.
- *The toolbox does not consider specific criteria within nature and form of explanations.* There is a need to consider specific criteria related to each answer format. Such criteria should be guided by the philosophical models that are presented earlier. A complete and meaningful K-12 schema should not merely list the types of answer formats (such as CDN, CIS, CDR, etc.). It should further explore the nature of the elements that make up these formats.

Table 2.2

The Pragmatic toolbox for scientific explanation as derived from Weber et al.'s (2013) toolbox

Question type	Answer/explanation formats	Philosophical Model(s)
Questions about plain facts: 'Why is it the case that A?' Example: Why do objects fall to the ground?	-Causal deductive nomological format (CDN): (1) includes an explicit relevant population, (2) must be purely inferential, and (3) highlights the practical usefulness of causal knowledge. Example: The causal relation between smoking and lung cancer. -Causal inductive statistical format (CIS): Similar to CDN by adding a causality requirement to the IS model. -Causal default rule (CDR): there is no requirement for deduction or induction with an exact probability value. Example: Builds usually fly; penguins do not fly.	Hempel (1965) believes that all explanation-seeking questions about particular facts are of this form. Kitcher (1989) states that an answer to the question 'Why is the case that S' is explain the conditions under which S occurred. Salmon (1998) always refers to the explanandum as "the fact-to-be explained"

(Table continues)

Table 2.2 (Continued)

Question type	Answer/explanation formats	Philosophical Model(s)
	-Positive causal relevant (PCR): necessitates that a valid explanation must contain at least one PCF; otherwise the explanation is untrue. -Positive and negative causal factors (PNC): both PCF and NCF are conditions for an adequate explanation.	
Contrastive questions: 'Why A rather than B' Example: Why is the Earth constantly changing rather than remaining the same?	-Reality to alternative format: Why A rather than B requires first describing A (what really happened), explaining why A happened, and then presenting B, an alternative scenario that if occurred would produce a different result. Example: Why does a plastic ruler (A) rather than a wooden rod (B) attract small pieces of paper? -Real contrast format: cases where we do not have one alternative but more. Example: Plastic ruler rubbed on hair attracting piece of paper rather than pieces of leather. Plastic ruler rather than wooden rod attracting pieces of paper, and so on.	Van Fraassen (1980) refutes questions of the form 'Why is it the case that A' and states that all scientific question must be, either explicitly or implicitly of contrastive form.
Resemblance questions: focus on similarities between events as opposed to only differences between events or on a single event. Example 'How are plants, animals and environments similar from current plants, animals and environments?'	Top-down unification format: events to be unified belong to the same law (or set of laws) of nature. Example: Events explained by Newton's laws of universal gravitation and Coulomb's law of electrostatics (set of two laws explaining several events). -Bottom-up unification format – is more concerned with societal behavior that are caused by common causal factor that do not necessarily derive from a law of nature.	Kitcher's (1989) unification model of explanation (with an added causality requirement)

Towards a Philosophically Guided Schema for Studying Students' Scientific Explanations

Guided by philosophical models of, and approaches to studying scientific explanation, this section discusses their usefulness in developing a K-12 science explanation schema. Table 2.1 summarizes different types of scientific explanations derived from multiple philosophical models of explanation. The table also includes relevant examples of each type of explanation, and summarizes some of the requirements of each of these types. All forms of explanations in

Table 2.1 take place within a pragmatic context. Table 2.2 summarizes answer formats from Weber et al.'s (2013) toolbox for examining scientific explanation.

PART 2: An Examination of Research On Scientific Explanation in Science Education

After over six decades of research in scientific explanation in philosophy of science, there is still no consensus regarding a favorite unified model of explanation (de Regt, 2009). Previous philosophical debates have resulted in a variety of models of scientific explanation and significant disagreements among philosophers of science. Clearly, the notions of scientific explanation and scientific understanding are complicated philosophical constructs. Nonetheless, scientific explanation in the science classroom has been of interest to researchers in science education for over 40 years. While the majority of studies aimed at assessing students' scientific explanations, some attempted to develop models that focus solely on explanations. An examination of the literature reveals that research related to scientific explanation in science education has been conducted along three distinct lines: (1) one that employed general approaches to all students' answers; (2) another that utilized models of argumentation to assess explanations; and (3) the third aimed at developing models specific to scientific explanation. Studies in each of these lines are presented in this section of the chapter.

Examination of *All* Students' Answers: General Approaches to Analyzing Scientific Explanation

As early as the 1990's it was evident that children's scientific explanations were important in examining how they learn about science and science processes (Metz, 1991). Research on explanations in science education, however, has shown that teachers, and in turn, students lack competence in constructing scientific explanation (e.g. Erduran, et al., 2004; Haefner & Zembal-Saul, 2004; Yao, et al., 2016). The need to examine different ways to improve students' explanation construction and teachers' explanation assessment became apparent. It was natural that researchers started with assessing students' explanations of scientific investigations.

There exists a consensus in science education literature on the importance of constructing scientific explanations. Some researchers argue that the ability to construct sound scientific explanations shows evidence of deep learning and understanding of scientific concepts (e.g., Sevian & Gonsalves, 2008). Other researchers assert that constructing scientific explanations can help students learn more about their own understanding (e.g., Colombo, 2017; Beyer & Davis, 2008). In addition, some researchers in science education believe that constructing explanations in science helps develop students' scientific literacy (e.g., PISA, 2015; Ryder, 2001).

While the goal of constructing scientific explanations is important in science, the problem with it is that teachers lack the knowledge and guidance to achieve it effectively (Beyer & Davis, 2008). As previously discussed, there is no detailed articulation of the nature of scientific explanation in science education. To this date, the lack of a clear conceptualization and articulation of the nature of scientific explanation has not been appropriately addressed in science education (Tang, 2016). Thus, the studies discussed in this line of research examined students' scientific explanations in various school settings using various general, at times incomplete, vague, or flawed, approaches.

As early as 1991, Kathleen Metz noted the importance of explanations in science and drew on the distinction between explanations and predictions. In her work, the researcher was interested in examining children's causal knowledge through the analysis of their scientific explanations. In particular, Metz focused on changes in content and form of the explanations that three to nine-year old children constructed in the context of movement and jamming in gears.

Thirty-two children were asked to answer physics-related questions about gear configurations. Questions in this study included predictions (e.g., “What would happen when you turned the knob?” p. 788), and explanations (e.g., “Why would [repeat of S’s prediction]”? p. 788). Following the child’s explanation, the interviewer asked the him/her to check their predictions. An examination of the study’s framework of analysis of children’s answers revealed a novel method of analyzing explanations from a cognitive perspective. In particular. Metz categorized explanations into (1) conceptual entities, (2) actions, and (3) relations. Conceptual entities, according to Metz, included objects that were directly related to the event-to-be-explained (such as the knob in a certain gear situation).

Based on the explanations obtained, Metz developed 11 explanation types that were used to categorize the explanations generated by the participants. Examples of these types included “Function of a circle”, “Function of a knob”, “Connections: gear-teeth”, etc. (All 11 explanation types can be found in Metz, 1991, p. 789). It was unclear how these types were developed as there was no theoretical framework that supported the analysis. Furthermore, these types seem to be specific only to the task conducted in this study. Thus, it was difficult to generalize this framework and use it to assess explanations in other contexts.

Many other researchers in science education were interested in examining students’ scientific explanations but considered *all* students’ answers to be explanations. Researchers in these studies did not make clear distinctions between explanations and other practices (such as descriptions, predictions, reasoning, and justifications). In many cases, different terms were used interchangeably (such as explanations and reasoning; explanations and ideas; etc.) that resulted in more ambiguity about the validity of the assessment process. See for example Forbes, Lange, Moller, Biggers, Laux, and Zangori (2014); Kesonen, Asikainen, and Hirvonen (2017);

Kokkonen and Mäntylä (2018); Mestad and Kolstø (2017); Meyer and Woodruff (1997); Lawson; Drake; Johnson; Kwon; Scarpone (2000); Peker and Wallace (2011); Southard, Espindola, Zaepfel, and Bolger (2017); Zangori, Forbes and Schwarz (2015). For instance, Kesonen, et al. (2017) used the terms explanations, reasoning and students' ideas interchangeably in their analysis of students' answers to questions about the behavior of light. Similarly, Mestad and Kolstø (2016) did not distinguish between explanations, descriptions and interpretations. Their analysis generated two type of explanations (event-focused explanations and object-focused explanations) that were based on their findings. Zangori et al. (2015) focused on model-based explanations, and Kokkonen and Mäntylä (2018) applied a concept of learning to examine changes in university students' explanations to questions related to DC circuits. These researchers' findingsfocused on the types of concepts students used in their answers. Similarly, Zuzovsky and Tamir (1999) and Forbes et al. (2014) examined all students' answers as explanations on internationally science assessment tests. They analyzed answers given by students to questions on the Third International Mathematics and Science Study (TIMSS); while Forbes et al. compared and analyzed fourth-grade science classrooms in Germany and the US. A number of other studies, based their analysis on previous theoretical work, but did not make clear the distinction between scientific explanations and other constructs. For example, Peker and Wallace (2011) examined high school students' explanations using an epistemological characterization of reasoning. Other researchers (e.g., Sandoval, 2003; De Vries, Lund, & Baker, 2002) examined all students' answers as explanations within a technologically-rich collaborative learning environment.

Another trend in this line of research focused on a special type of explanations, mainly teleological explanations. Teleological explanations, abundant in biology, are statements that involve explaining phenomena in terms of purposes, functions and goals (e.g., Trommler, Gresch, & Hammann, 2018). Stemming from the researchers' belief that teleological bias hinders learning, in a study on teleological explanations, Trommler et al. explored the reasons why students prefer teleological explanations. Other studies on teleological explanations included the work of Halls, Ainsworth, and Oliver (2018); Kampourakis. Pavlidi, and Palaiokrassa (2012); Talanquer (2007); Tamir & Zohar (1991); Kampourakis, Silveira, and Strasser (2016).

Aside from studies on teleological explanations, which constitute a specific type of research that focuses on biology education, other general and unsupported approaches to examining students' scientific explanations resulted in a dearth of clear conceptualization of scientific explanation. In the absence of these guidelines, research on scientific explanation has resorted to models on peripheral topics, such as Toulmin's model of argumentation, to assess explanations without necessarily making a convincing case that arguments are some type of explanations (e.g., Delen & Krajcik 2018; Peker & Wallace 2011; Yang & Wang 2014). The following section provides a brief overview of the studies that fall in the second line of research focusing on the ways by which researchers justified, if any, such use.

Using of Models of Argumentation to Assess Students' Explanations

The notions of scientific argumentation and explanation in science have been of interest to researchers in science education for decades. Some researchers in science education assert that while they are not the same, explanation and argumentation are complementary (e.g., Berland & Reiser, 2009). Other researchers consider them to belong to the same scientific practice (e.g., McNeill and Krajcik, 2008).

One of the major reasons for adopting models of argumentation to assess scientific explanations, researchers in this line of research argued, was due to the need for a simple and practical framework that provided an account for teachers of what a *good* explanation looks like. See, for example, Berland and Reiser (2009); Bell and Linn (2000); Driver, Newton, and Osborne (2000); Erduran et al. (2004); Forbes et al.(2014); McNeill and Krajcik (2008); McNeill, Lizotte, Krajcik, and Marx (2009); Ruiz-Primo Tsai, and Schneider (2010); Sandoval, (2003); Wang (2014). In these studies, the researchers employed a version of Toulmin's model of argumentation, usually known as the Claim-Evidence-Reasoning (CER) model, to examine students' explanations in various settings. It is important here to note that the CER model was originally developed to examine arguments, not explanations, in science. According to Toulmin's original model (1958), the main goal of a scientific argument is to determine the related characteristics of a claim. Thus, Toulmin developed an argument structure that included a claim, evidence or data that support or oppose the claim, and assumptions (backing) and principles (warrants) on which the claim is based.

Science education researchers who believe that arguments and explanations are different still share different views on the employment of each construct in teaching and learning. For example, Bell and Linn (2000) considered argument construction as a priority. They asserted that explanation is a by-product of argument construction (also see Osborne, Erduran & Simon, 2004). Other researchers believed that the focus was explanation construction, and that good explanations resulted in good arguments (e.g., McNeill, Lizotte, Krajcik, & Marx, 2006; Sandoval & Reiser, 2004). However, in their analysis of students' explanations these researchers resorted to the structure of a scientific argument (claim-evidence-reasoning) in order to aid students' construction of scientific explanations.

On the other hand, researchers who considered explanations and arguments as synonymous adopted the CER model and argued that the term “scientific explanation” entailed pedagogical goals of both argumentation and explanation (McNeill & Krajcik, 2008). Others examined students’ explanations as arguments within the approach of constructing and defending scientific explanations (Berland & Resier, 2009). In their work, McNeill and Krajcik explicitly stated that they used the term ‘explanation’ in order to match their work with the national and state standards that teachers require to attain the state goal in helping “students construct scientific explanations about phenomena where they justify their claims using appropriate evidence and scientific principles” (p. 54). In their discussion about explanations, the researchers defined an explanation as a linguistic construct that included a claim (i.e. a conclusion of a problem); evidence (data that support the claim); and reasoning (a justification for why the evidence supports the claim). Drawing on the work of McNeill and Krajick (2008), Ruiz-Primo et al. (2010) asserted that “scientific explanations should connect patterns of data with claims about what the data mean” (p. 586).

Other studies implemented the CER framework incorporating it with other aspects such as technology-enhanced environments, scaffold-based environments, a focus on written explanations, etc.. In this regard, Delen and Krajcik (2018) used the CER framework to assess students’ scientific explanations after using a new mobile application that helped students collect and use data to construct explanations. Jang and Hand (2016) used a model of argumentation to assess students’ argumentative and explanatory writings. In their study, they examined the value of using a scaffolded critique framework to aid students in the construction of explanations and arguments in science. While the researchers distinguished between two types of writing: argumentative writing and explanatory writing, their developed Scaffolded Critique Framework (SCF) was based on two criteria: “one is going back to examine the alignment between their [the students'] claims and evidence and their readings and the second is going forward to examine if these claims and evidence and comparisons are coherent" (p. 1220).

In a number of studies in this line of research, researchers used the concepts of scientific explanation and argumentation interchangeably without making a convincing case for such use. For example, Berland and Resier (2009) first recognized that arguments are different from explanations and indicated that they are complimentary. They further argued that science teachers should be provided with appropriate instructional guidelines that focus on the distinction between explanations and arguments in science. However, later in their paper the researchers used the two terms interchangeably when they regarded the claim "some birds survived because they ate a specific plant" as an explanation because it included the three elements (claim, evidence, and reasoning). What is more, in another study by Osborne and Patterson (2010), the researchers asserted that the Berland's and Resier's statement should not have been regarded as an explanation. Osborne and Patterson, however, stated that the statement "some birds survived" was not actually a claim but rather a statement of a fact. In addition, they added that the statement "they ate a specific plant" was neither data nor evidence but rather a description. Nonetheless, it seems that the problem that Osborne and Patterson had with Berland and Resier's analysis was not in the fact that they used the CER framework to examine explanations, but in how they used it. In another study on explanation, Sandoval and Millwood (2005) analyzed the quality of the use of evidence among high school students' scientific explanations of questions about natural selection. The researchers used a strategy that "assess[es] the warrant of explanatory claims, the sufficiency of the evidence explicitly cited for claims, and students' rhetorical use of specific inscriptions in their arguments" (p. 23).

Evidently, there is an emergent confusion in the field of science education regarding the use of the terms “argument” and “explanation”. Osborne and Patterson (2011) sought to clarify such confusion by distinguishing between the two terms. The researchers argued that a necessary distinction is important in identifying the nature of the activity that takes place in a science classroom. Thus, they discussed criteria of both scientific argumentation and explanation supported by empirical and philosophical work. They also provided definitions of both terms that highlighted the differences between the two constructs. In particular, Osborne and Patterson believed that arguing and explaining are two different linguistic processes with different epistemic functions: while explanations aim to provide understanding, arguments aim to convince. Another fundamental difference that the researchers discussed was related to the degree of tentativeness of the phenomenon at hand. According to Osborne and Patterson, while explaining a phenomenon, the phenomenon itself is more or less taken for granted. However, while constructing an argument, the degree of certainty of the phenomenon is the focus of the argument. Along the same lines, Brigandt (2016) contributed to the debate on whether or not scientific explanations should be distinguished from scientific arguments. In his article, Brigandt asserted that a distinction between the two terms is important to science education. More precisely, he argued that since explanations and arguments have different epistemic goals, they should also have different standards of adequacy. In order to understand what counts as a *good* explanation, Brigandt continued, science educators should focus on explanatory adequacy rather than focusing on evidence-based argumentation. The latter focus could “obscure such standards of what makes an explanation explanatory” (p. 251). Finally, Tang (2016) provided further insight on the need to end the conflation between explanations and arguments in the science classroom. Tang asserted that while Toulmin’s model of argumentation, or the Claim-Evidence-Reasoning model, is appropriate for assessing arguments that result from empirical work, it is not suitable for assessing theoretical-based explanations that provide causal accounts of natural phenomena.

Philosophers of science have also debated whether or not explanations and arguments are the same practice. While some philosophers regarded explanations as arguments (e.g., Hausman, 1998; Hempel, 1965), others discussed their differences. In fact, the assertion in philosophy that explanations are arguments faced a lot of criticism. One of the most common critiques of Hempel's view of explanations as arguments is related to the problem of irrelevant premises. Because Hempel considered explanations as arguments, superfluous premises influenced whether the statements he claimed to be explanations were in fact explanatory. More specifically, philosophers argued that if one adds more premises to an argument, it will still be an argument. However, adding more (irrelevant) premises to an explanation does not preserve the explanatory power of the original explanation (e.g., Weber et al., 2013).

A closer examination of the differences between scientific explanations and arguments from a philosophical perspective shows that while arguments request evidence to persuade the audience, explanations require causes and/or references to theories, laws and regularities. Additionally, in everyday life, the meaning of argument is different from that used in the context of science teaching and learning. For instance, most people consider arguments as a form of dispute. However, arguments in science are statements in which evidence is provided to support a certain claim (e.g., Mayes, 2010). Similarly, explanations in everyday situations are usually thought of as descriptions or accounts of the facts. In science, however, explanations are answers to *why*-questions in reference to causal links, statistical-probabilistic laws, and/or natural regularities. An easy way to distinguish between them, Mayes suggested, is to remember that an argument answers the question *How do you know?* whereas an explanation answers the question

Why is it so?

Summary. Research on scientific argumentation and scientific explanation has been of interest to researchers in science education for decades. While there has been conceptual work done on the process of assessing scientific argumentation, the case is not the same for the assessment of scientific explanation. Thus, some researchers in science education have resorted to frameworks of argumentation to assess explanations. While this approach seems to be accepted among a considerable number of researchers in science education, other researchers have made it clear that the two notions are distinct with different goals. These researchers have called for an assessment tool that is unique for scientific explanation.

Researchers who advocate the need for a framework unique to scientific explanation in science education belong to the third line of research. Studies in this line of research include the development of frameworks from different approaches. Thus, the third line of research is of most interest to this study. In the next section, a critical examination of these attempts in science education is presented.

Development of Models Specific to Scientific Explanation

As early as 1970, researchers within this line of research attempted to develop models that aimed to provide guidelines to the process of explanation construction. This was due to the lack of a clear conceptualization of explanation in science education and the conflation between explanations on the one hand and other concepts such as descriptions, predictions, and arguments on the other. While these models are important attempts for conceptualizing the process of constructing adequate scientific explanations, they still had some gaps. Some of the gaps included lack of clear distinctions between explanations and non-explanations (such as descriptions and predictions), misinterpretation or misrepresentation of philosophical models of explanation, among other issues. It is worth mentioning here that a lot of these studies pre-date the development of the NGSS, which gave more attention to the role of scientific explanation in the science classroom. The following section presents a critical review of the various attempts undertaken by researchers in science education to develop models and guidelines that are specific to scientific explanation.

Among the first attempts to explore the nature of scientific explanation within the context of the science classroom included the work of a number of researchers in science education such as Smith and Meux (1970), Ivany and Oguntonade (1972), and Dagher and Cossman (1992). Smith and Meux classified explanations into different types by analyzing classroom transcripts from a pedagogical viewpoint. They categorized explanations into nine types: normative explanations, empirical-subsumptive explanations, judgmental explanations, procedural explanations, sequent explanations, teleological explanations, explanations by consequences, and mechanical explanations. These categories covered a wide range of science skills. For example, normative explanation was based on providing evidence that the phenomenon-to-be-explained was subsumed under some rule or norm. In addition, empirical-subsumptive explanations were subsumed under empirical generalization of law, judgmental explanations were based on highly probable generalizations or laws. Procedural explanations included statements of a series of actions that led to the phenomenon-to-be-explained, and so on.

Along the same lines of research, Ivany and Oguntonade (1972) examined physics teachers' explanations and matched them with a set of pre-existing criteria using a philosophical framework. Based on their analysis of teachers' explanations transcripts, the researchers listed a set of criteria of explanations: universal laws, constructs, analogies, and historical accounts. A verbal explanation analysis instrument was then developed in order to analyze these transcripts. The generated list of criteria offered guidelines to meaningful explanation construction, and the results of the study were, in fact, useful in highlighting the most prevalent types of explanations. Frequency counts of verbal explanations revealed that constructs were the most prevalent in teachers' explanations. In addition, the researchers concluded that teacher lecturing was the most predominant mode of verbal explanation. Therefore, it was recommended that teachers needed training in purposeful ways to teach meaningful construction of verbal explanations, the use of appropriate explanations, and the adequate understanding of historical accounts of scientific investigations.

Dagher and Cossman's (1992) approach was different from the above two studies in that "[they] sought guidance for what constitutes an explanation from philosophy of science" (p. 362). In this study, the researchers analyzed explanations produced by science teachers in their classrooms. The analysis resulted in the generation of ten types of explanations: analogical, anthropomorphic, functional, genetic, mechanical, metaphysical, practical, rational, tautological, and teleological. A detailed account of the ten types of explanation can be found in the original paper (pp. 364-366). These ten types were then subsumed under literature-based categories. The researchers asserted that literature from philosophy of science guided the category formation procedure. Thus, Dagher and Cossman constructed a framework that portrayed the relationship between the subsuming categories (from philosophy of science) and the generated types (based on their analysis in their study).

In their discussion of the philosophy literature related to their generated types, the researchers argued that theoretical explanations were the most prevalent type of explanation on which they based their generated list. A theoretical explanation "rationalizes facts and render them intelligible to a mind seeking to understand". A theoretical explanation was then divided into two categories: genuine explanations that are either true or false and spurious explanations that cannot be falsified. A fourth category that fell out of the theoretical explanation category was the practical (how-to) category that was left undefined in this paper.

The three previous studies were the earliest investigations in science education that focused on the role of meaningful explanation construction in the science classroom. Furthermore, researchers in these studies asserted that meaningful construction of explanation was neglected. So it seems that since the 1970's researchers in science education have been calling for improving teachers' education program in this regard (e.g., Ivany & Oguntonade, 1972), and have attempted to develop useful tools for teachers that aid them in meaningful explanation construction. These studies were promising, but the models generated were too broad to be used as pedagogical guidelines. Research on developing such tools is still ongoing; it has progressed to include deeper conceptualization of philosophical models of explanations.

Braaten and Windschitl were two of the first researchers to develop a tool specific to scientific explanations that sought to aid science teachers in making the practice happen and offer philosophical support for this tool. Braaten and Windschitl (2010) asserted that in order for science teachers to be able to promote students' construction of explanation in science, science educators must provide them with more guidance about the nature of scientific explanation. The researchers based their conceptualization of explanation on their interpretations of philosophical models that examined the structure and role of scientific explanations. The authors considered the philosophical body of work as a tool for analyzing how explanations are generated in the science classroom and how researchers in science education could design learning environments to promote meaningful construction of explanations within the science classroom.

One of the purposes of Braaten and Windschitl's (2010) study was to develop an explanation tool that helped to "organize teachers' thoughts about scientific explanations in their classrooms" (p. 661). Before presenting their Explanation Tool, the authors provided a brief review of each of the five philosophical models of scientific explanation: The Deductive-Nomological Model, the Statistical-Probabilistic Model, the Causal Model, the Pragmatic Model, and the Unification Model of Explanation.

Following their review of the philosophical literature of scientific explanation and its application to the science classroom, Braaten and Windschitl briefly discussed the ways in which explanation is portrayed in science education research, and how science education reform documents defined scientific explanation. They concluded that there was some ambiguity in examining students' explanations within the science classroom. In particular, the researchers found that in most of the studies on scientific explanation in science education students were directed to use evidence and reasoning to support their answers resulting in statements of justified belief or argument, but not necessarily a scientific explanation. Furthermore, the researchers' examination of the reform documents revealed that these documents call for an emphasis on scientific explanations pushing teachers away from focusing on solely describing, measuring, and observing and focusing more on explaining and understanding. However, these documents, the researchers revealed, did not offer specific guidelines that help teachers design learning environments that promoted students' scientific explanation. Hence, Braaten and Windschitl sought to fill this gap by "developing conceptual and pedagogical tools offering heuristic value for teachers to carry out specific instructional practices pressing for the co-constructions of scientific explanations in science classrooms" (p. 657).

The development of this pedagogical tool was the result of the researchers' work with science teachers on scientific explanation since 2007. In the process of developing the tool, the researchers first characterized the attributes of what they considered to be a good scientific explanation. Their rubric, the Explanation Tool, was developed in such a way that it (1) utilized major scientific theories, (2) sought underlying theoretical causes for observable natural phenomena, and (3) when applicable, employed mathematical models to depict patterns in data. Additionally, Braaten and Windschitl made a powerful decision in using a variety of models of scientific explanation to portray the actual practice of science.

Examining the Explanation Tool (in Table 3, p. 662), however, reveals some discrepancies between what the researchers had been theorizing about a conceptual tool for assessing scientific explanation, and how the tool actually depicted scientific explanation. First, the continuum-nature of the rubric referred to the 'depth' of students' explanation – a quality that was left undefined in their article. The rubric was two dimensional: on one dimension was the discrete distinction between explanations with theoretical component vs. explanations with mathematical component, and the other continuous increasing-in-depth dimension was the *What*, the *How*, and the *Why*.

First, even though throughout their paper Braaten and Windschitl (2010) emphasized the importance of distinguishing between descriptions and explanations, one of the three depth-levels of an explanation in their explanation tool was the *What* level. In particular, the researchers considered the act of describing what happens as an explanation – a first depth level of an explanation. They further elaborated that within the What-explanation "student describes, summarizes, or restates a pattern or trend in data without making a connection to any unobservable/theoretical components" (p. 662). While descriptions are essential in science, they are not types of explanations. Developing a tool that accurately examines scientific explanations should make a clear distinction between descriptions and explanations. Such a tool should also draw on the importance of the role of a description as a part rather than a type of explanation.

The second depth level (out of three levels) in the Explanation Tool was the *How* in which students describe “how or partial why something happened. Students address unobservable/theoretical components tangentially” (p. 662). It remained unclear what a “partial why” meant and the ways by which it was related to the how. In addition, throughout their paper, Braaten and Windschitl did not discuss how tangentially addressing theoretical components provided some kind of explanation.

In the last and highest depth-level, the *Why* “explains why something happened” or “why a mathematical model accounts for a phenomenon” (p. 662). However, the assertion: “to explain something, a student explains why something happened” made by Braaten and Windschitl does not accurately capture what a deep level of an explanation entails. Furthermore, the researchers’ Explanation Tool did not offer clear guidelines for teachers to help students construct meaningful explanations. In fact, Braaten and Windschitl stated that they purposefully developed an “oversimplified framework for thinking about scientific explanations” (p. 663). It remained unclear how the teachers utilized the tool to rephrase the science questions asked in the classroom or to improve their instructional materials to promote explanatory reasoning, as claimed by the researchers.

Braaten and Windschitl (2010) were successful in shedding light on the problem of the lack of meaningful explanation construction in the science classroom, in science education literature, and in science curricula. In addition, they made an informed decision by tackling this problem from a philosophical perspective. However, their claims and their interpretations of the philosophical models of explanation were not reflected in the development of their explanation tool, which did offer a solution to the problem of the lack of guidelines to meaningful explanation construction in science education.

During the past four years and following the work of Braaten and Windschitl (2010) there have been a few attempts for developing new models to analyze scientific explanations. Yao et al. (2016) work focused on developing a framework based on philosophical models of explanations for examining students' scientific explanations. The Phenomenon-Theory-Data-Reasoning (PTDR) framework was developed as a tool for teachers to aid in the instruction about scientific explanations. In addition, De Andrade, Freire, and Baptista (2017) presented a system of analysis of students' explanations in science. They adopted the causal and unification models of explanations to guide their analysis. Finally, the most recent work in this line of research is Papadouris, Vokos, and Constantinou's (2018) work on the pursuit of a better explanation through the development of a framework for science teaching and learning. A critical examination of these three studies is provided in what follows.

In addition to philosophical models of explanation, Yao et al.'s (2016) PTDR framework was developed based on students' different content characteristics, diverse learning backgrounds, and their features. The researchers interpreted several philosophical models of explanation including Hempel's covering law model, Salmon's causal model, the unificationist account of explanation, and the pragmatic theory of explanation. In developing their framework, and unlike many researchers in science education (e.g., McNeill, Lizotte, Krajcik, & Marx, 2006; Songer, Kelcey, & Gotwals, 2009), Yao et al. (2016) asserted that an educational framework of scientific explanation ought to be different from that of scientific argumentation. In particular, and due to the existence of sound philosophical support, the researcher argued that an educational framework of scientific explanation should be developed based on philosophical models of scientific explanation.

Following a brief discussion of the philosophical models involved in the construction of the PTDR framework, the researchers proposed their hypothesis for a learning progression of scientific explanation. First, the researchers adopted the syntax structure of a scientific explanation from Hempel's covering law model. Thus, they suggested that an explanandum, general laws, and antecedent conditions constitute three key components of a scientific explanation. In addition to the syntax structure, Yao et al. (2016) incorporated the causal-mechanical model and focused on the importance of causal interactions and causal mechanisms. Thus, they added *reasoning* as a fourth component in their analysis.

Yao et al. (2016) asserted that when constructing a scientific explanation, a PTDR framework examines the identification of the phenomenon (P), the theories (T) and data (D) used to explain the phenomenon, and the association made among the data, theories and the phenomenon-to-be-explained (i.e. Reasoning, R). A few important issues should be discussed regarding the mapping of philosophical theories into the development of the PTDR. First, there is an important philosophical distinction between explanandum and phenomenon that Yao et al. did not consider. According to the covering law model, an explanandum is a statement (or a group of statements) of a phenomenon rather than the phenomenon itself. A natural phenomenon exists independent of people's statements about it. Furthermore, statements about phenomena (i.e. explananda) change as people gain more knowledge without necessarily any change in the phenomenon itself. Therefore, precision is required when referring to explanandum as statements about phenomena rather than phenomena.

Another issue with the PTDR framework is related to the *Reasoning* component and its projection from the causal-mechanical model. A more logical mapping of the philosophical model of causality could have simply been causal links, causal interactions or causal mechanism instead of reasoning. It was unclear how *reasoning* was a part of a scientific explanation. Thus, the process of reasoning should be clarified in terms of what counts as a process of reasoning, reasoning patterns included in the framework, and the distinction and relationship between Theory and Reasoning.

Yao et al.'s (2016) learning progression of scientific explanation was based on the PTDR framework and suggested that each component in the PTDR framework is comprised of two levels, a lower and an upper level, with some components having three levels (an intermediate level). For example, a lower level Phenomenon component describes a phenomenon in a clear manner and includes a simple relationship between variables. On the other hand, an upper level Phenomenon component represents the phenomenon from within a real context and includes multiple variables connected in a complex relationship. Another example is that of the Theory component where a low level includes the application of scientific ideas and law-like statements with teachers' guidance or with the help of instructional materials. An intermediate level of the Theory component emphasizes the use of general laws independently, and an upper level include the independent selection of the laws and theories and linking them systematically with the context. However, the criterion of guidance (whether from teachers or instructional materials) in the Theory component seems to be problematic in that such a framework does not help students construct scientific explanations, but is used to analyze their explanations during classroom observations.

While some researchers in science education attempted to develop frameworks to analyze students' scientific explanations, others' frameworks focused on certain kinds of explanations or specific issues related to explanations. Research in science education reveals several frameworks that target scientific explanations, with various foci as well as different perspectives on what counts as a good explanation. A brief examination of these studies is presented in the following section.

More recently, de Andrade et al. (2017) developed a framework of what counts as a *good* explanation and proposed a system of analysis to categorize students' scientific explanations based on the framework they developed. The system, the researchers argued, was "conceptualized and developed based on theories and models of scientific explanations, science education literature, and from examples of students' explanations collected by an open-ended questionnaire" (p. 1). Using answers from an open-ended questionnaire, the researchers categorized students' answers and developed a system to analyze their explanations.

The study started with a definition of scientific explanation and a discussion of the key characteristics of a 'good' explanation in science. Among the various philosophical models of explanations, the researchers stated that they adopted two models that they considered to be relevant to their framework - the causal and the unification models of explanation – to guide their analysis. In defining a scientific explanation for science education, de Andrade et al. (2017) listed a few of its characteristics, such as the fact that explanations in science are "more systematic, deeper, and more accurate than common sense explanation" (p. 3). In addition, the researchers asserted that scientific explanations explain a new phenomenon by referring it to other scientific facts and theories and/or by identifying its causes. Following the brief definition, the researchers based their assessment of a good explanation on the work of Braaten and

Windschitl's (2011) framework. Thus, the characteristics of a good scientific explanation, according to de Andrade et al. (2017), included (1) relevance, (2) conceptual framework based on theoretical ideas of science, (3) a trace of the causal story of the phenomenon-to-be-explained, and (4) an appropriate level of representation.

Based on the above definition and key characteristics of a scientific explanation, the researchers conducted an empirical study in order to examine how students constructed scientific explanations about chemical phenomena. In addition, de Andrade et al. (2017) aimed at exploring the ways by which students can be supported to improve their explanation construction practice. Students were asked to generate explanations of natural phenomena, their answers were analyzed, and the quality of their explanations was examined. The questionnaire included questions about four phenomena to be explained. The phenomena included (1) mixing liquids with different densities, (2) dissolving sugar in water, (3) water condensation on the surface of a cold can, and (4) thermal expansion of a gas. Before completing the questionnaire, the teachers discussed with the students what explanation, description, and justification meant. The teachers then constructed a scientific explanation of the phenomenon of the diffusion of floral oil scent around a room as an exemplar, and the students were shown how to describe observable features, how to include predictions, and how to identify scientific ideas related to the phenomenon. Details regarding this process were not mentioned in the paper, however. The researchers then described the process by which they developed a system of analysis of explanations and how this system was utilized to categorize students' answers.

In an aim to analyze students' scientific explanations, the researchers "outlined a hypothetical good scientific explanation, considering students' curricular level and the framework of scientific explanation previously developed" (p. 8) presenting a flowchart to illustrate the questions of the scientific concepts involved, causal links, and representational levels involved. An outline of a good explanation of a phenomenon, according to the figure presented (Figure 1, p. 9), included three boxes. The first box targets *What happened*. Through this box, two arrows emerge, both of which ask *How it happens*. Each arrow leads to two levels: the submicroscopic level and the macroscopic level. The two levels are then connected with a two-directional dotted arrow of *Why it happens.* For example, according to the figure, an outline of a good explanation of the second question included a statement of what happened ("Sugar was added to a glass of water. As a result: It is not possible to distinguish a glass of water from the other", p. 9), a submicroscopic statement about how it happens ("In dissolving, water manages to separate sugar particles; and as the sugar particles move in between the water particles, the water particles move in between the sugar", p.9), and a macroscopic statement ("A physical phenomenon happens: dissolving – the dissolution of sugar (solute) in the water (solvent). It leads to the formation of a single medium or a homogenous solution, colourless and odorless solution assuming the case of complete dissolution", p. 9). The two-directional dotted arrow between the submicroscopic and the macroscopic levels remained unclear.

Based on the key characteristics of a good explanation, students' answers were then categorized as non-explanations, pseudo-explanations, and explanations. Non-explanations were answers to questions that (1) did not include relevant information to the phenomenon, (2) did not include conceptual framework, (3) included restatements of what was previously presented, (4) did not provide additional insights, and (5) did not enable understanding of the phenomenon. Pseudo-explanations included answers in which students described what happened but "paid little attention to the specific entities and the underlying processes that produced the phenomenon and that presented a poor causal scheme" (p. 10). Finally, explanations were answers that included how and why the phenomenon occurred and "presented logical and coherent causal stories that relied on a conceptual framework, in which observed events are attributed to the underlying processes" (p. 10).

While de Andrade et al.'s (2017) system of analysis of an explanation provided important criteria, such as the what, the how and the why, the outline and the key characteristics are content-specific and non-generalizable. In addition, the researchers' choice of the questions facilitated both the submicroscopic and the macroscopic levels. The same framework is not applicable – as is- in all science-related topics, however. In addition, the researchers added causality as one of the criteria without taking into consideration that not all scientific phenomena are causal. Furthermore, it was unclear how some of these criteria were measured. For example, it was unclear how one can analyze a statement in terms of whether or not it attains understanding of the phenomenon.

After categorizing students' answers into the previous three categories, a closer examination of the answers was conducted in which the researchers examined particular criteria of each category. This process resulted in an analysis of the nature of causal relations. Four characteristics of students' causal explanations emerged: (1) descriptions of events in terms of patterns and surface features without referencing the process involved, (2) associations of not-so-closely-connected events, (3) simple causal stories, and (4) complex causal stories.

Thus, a system of analysis was constructed based on the new categories of *pseudo-explanations* and *explanations*. Pseudo-explanations category was further divided into two categories: the descriptive explanations (that included two subcategories: macro and mix descriptions), and the associative explanations. Explanation category was further divided into two subcategories: simple and complex explanations. Associative explanations were defined as those that "associate pieces of information yet fail to establish how the information is related" (de Andrade et al., 2017, p. 14). Figure 2 presented a flowchart of the system of analysis for categorizing students' answers.

De Andrade et al.'s (2017) study provides a new insight to students' examining scientific explanations. However, a more general and non-content dependent framework is needed. Such a framework should extend beyond only two models of scientific explanation and should consider explanations that account for general laws and natural regularities, probabilistic laws and causal mechanisms.

A more recent work in this line of research is Papadouris et al.'s (2018) work on the pursuit of a better explanation through the development of a framework for science teaching and learning. The researchers in this study developed a theoretical account that focused on epistemic features of explanation as criteria for determining what constituted a *good* explanation. In this account, three features of explanations emerged: empirical validity, interpretive power, and generalizability. The researchers considered these features as a set of criteria that aim to help students in the construction and the evaluation of explanations. Empirical validity examined the extent to which predictions derived from an explanation align with the evidence at hand; interpretive power was concerned with the extent to which an explanation can account for the unfolding of the phenomenon; and generalizability focused on the extent of an explanation to offer a unifying framework that can explain other phenomena. Though not explicitly stated, it was clear that these three features were based on previous philosophical models of explanation.

While this study did not actually offer a practical framework that can be used by teachers and/or students to construct or assess scientific explanations, it contributed to the theoretical understanding of the importance of scientific explanation in the science classroom. Papadouris et al. (2018) discussed important epistemic features related to explanations that addressed the understanding of the nature of science. For example, the researchers called for appreciating that a phenomenon can be accounted for by more than one (rival) explanation, commit to the importance of empirical data, etc.

In addition to the previous studies examined thus far, this line of research also included studies in which researchers proposed various ways to assess scientific explanations that were not necessarily based on philosophical or theoretical background. While these studies offered empirical insight to the field of explanation construction, they were theoretically unsupported. A brief critical examination of these studies (e.g., Gilbert, Boulter, & Rutherford, 1998; Norris, Guilbert, Smith, Hakimelahi, & Philips, 2005; Yeo & Gilbert) is presented.

During the late 1990's Gilbert et al. (1998) sought to identify issues related to the role of models in scientific explanations. In so doing, they developed a typology of explanation that aimed at assessing the quality (or appropriateness) of scientific explanations. The researchers' typology addressed answers to questions regarding the phenomenon-to-be-explained. These questions included: "How does the phenomenon behave?", "Of what is the phenomenon composed?", "Why does the phenomenon behave as it does?", and "How might it behave under other conditions?" (pp. 85-87). In addition, the researchers categorized scientific explanations as appropriate and inappropriate. "[A]n appropriate explanation is one which adequately meets the needs of the questioner at the time that a question is asked. An appropriate explanation should facilitate and suggest directions for, as opposed to inhibiting subsequent questioning". (p. 87). They further listed four criteria that originated in the work of Toulmin (1972) to judge the value of an explanation: plausibility, parsimony, generalizability, and fruitfulness. On the other hand, "an inappropriate explanation is one where the match with experimentally derived data is not 'close' in terms of the application of these judgmental criteria by the inquirer" (p. 88).

Gilbert et al.'s (1998) typology of scientific explanations constituted a promising step towards a practical tool that could be used to examine students' scientific explanations. While the researchers did not reference their work on philosophical literature on explanation, their criteria were related to philosophical models of explanations. However, these criteria were too broad; more specific guidelines were still needed. For example, answering the question "Why does the phenomenon behave as it does" could result in various types of explanations from a philosophical perspective.

In another study, Norris et al. (2005) developed a framework to analyze "narrative explanations" in science. Their work aimed at the explanatory role of narratives. So, while their focus was more on narratives in science, they still tackled the notion of scientific explanation. First, Norris et al. (2005) defined narratives as verbally telling someone that something happened. The researchers then discussed various theoretical ideas associated with the elements of narratives, such as the narrator, the event, and the time. They also discussed the structure, agency, and purpose of narratives in general and in science. Norris et al.'s main purpose of the study was to evaluate narrative explanations in science. Thus, they briefly discussed the notion of explanation citing several researchers in science education, psychology, and philosophy of science. Acknowledging the difficulty of generating a complete and conclusive definition of scientific explanations, Norris et al. shed light on two main criteria of explanations: functions and types of explanations. In their work, they considered four types of explanations: deductive, probabilistic, functional, and genetic.

Norris et al. (2005) provided a summary of mainly the Deductive-Nomological (DN) model and discussed some of the problems related to it. They then discussed other types of explanations. They listed the functions and types of explanations and their characteristics (See Table 2, p. 550) arguing that function and type were not “always clearly separable” (p. 549). Thus, Table 2 in Norris et al.’s paper listed ten functions (or types) of explanations: interpretive, justificatory, descriptive, causal, deductive-nomological, statistical, functional, unification, pragmatic, and narrative explanations.

An examination of the characteristics of the different types of explanations revealed some ambiguities and inconsistencies. First, it was unclear whether Table 2 targeted scientific or non-scientific explanations. While deductive nomological, statistical and causal explanations were related to scientific explanations, interpretive and justificatory explanations were not necessarily so. An interpretive explanation, as defined by the researchers in Table 2, clarifies meaning; defines terms, propositions, or treatises; and assigns, develops, or expands meaning. This was clearly a general explanation and not a scientific explanation that aimed at attaining scientific understanding. Another problem with this list was the redundant definitions of several types of explanations. For example, Norris et al. (2005) stated that a justificatory explanation “explains by justifying why something was done”, or that a causal explanation “explains by citing a cause for events or laws”. There was no elaboration on the ways by which one justifies why something was done for example, or on the nature of causes of events or laws. Finally, the list in this table regarded description as explanations, denoted by “descriptive explanation”. A descriptive explanation, according to the researchers, explains by describing a process or structure. A clear distinction between explanations on the one hand and descriptions, interpretations, and justifications, on the other hand, was needed. Norris et al.’s (2005) main purpose was to highlight the importance of narratives in science education. Examining explanations was one phase in analyzing science narratives.

Along the same lines of examining language in the teaching of science, Yeo and Gilbert (2014) aimed at identifying the competencies that students needed in order to construct appropriate scientific explanations through developing a narrative account of explanation. The researchers examined explanations in three ways: in terms of their function, form, and level and generated a typology of scientific explanation. The typology included six types of explanation, their purposes, and the kind of question each type answered. Each of these types of explanation served different functions and implied that different explanations can be generated for a given phenomenon. Yeo and Gilbert regarded description and prediction as two types of scientific explanations. Contextualizing, intentional, interpretive and causal explanations were the other four types of explanation. The purpose of contextualizing, the researchers elaborated, was to give a phenomenon a name, an identity and enable it to be treated linguistically as a noun; thus answering the question "What exactly is being investigated?" (p. 1904).

In generating criteria about the quality of a scientific explanation, Yeo and Gilbert (2014) took a case study approach and used Lemke's multimodal framework to analyze scientific explanations within the context of science. Thus, the analysis was from a linguistic perspective and focused on how the language of science played a role in explanation construction. In addition to function and form, Yeo and Gilbert listed three levels of a given scientific explanation in terms of precision, abstraction, and complexity. Precision of an explanation was related to "its position in the evolution of research into a given phenomenon" (p. 1905); abstraction was another level of explanation that was related to the "process of simplification in which some aspects of an entity have been omitted or left unclear" (p. 1906); and finally complexity was the “measure of the composition/intricacy of an explanation” (p. 1908). In addition to the three levels of explanations, Yeo and Gilbert drew upon previous work on the levels of visualizations and adopted them for their assessment of scientific explanations. In particular, the researchers described the abstractness of explanations over three levels of visualizations: macro, sub-micro, and symbolic.

Yeo and Gilbert (2014) developed a multidimensional framework that was then used to analyze students’ explanations within the context of science. It offered a detailed analysis of students’ explanations and allowed for the evaluation of students’ answers using a new lens – a linguistic one - that did not focus solely on canonical correctness. However, the framework seemed to be more useful in exploring students’ ideas about a certain phenomenon rather than help them construct meaningful explanations about the phenomenon.

Other researchers attempted to develop models tackling various aspects of explanation construction. For example, Parnafes (2012) presented a theoretical model of the process of constructing scientific explanations using visual representations. The model was used to analyze students’ visual representations (diagrams, sketches, and general drawings) to explain the phases of the moon. Parnafes then developed a model to analyze these representations. In addition, Yang and Wang (2013) developed a teaching model that aimed at improving students’ explanation writing in science. The developed model (DCI) integrated “Descriptive explanation writing activity, Concept mapping, and an Interpretive explanation writing activity” (p. 531).

Another stance of developing models of explanation construction deals with science teaching explanations rather than scientists’ explanations. In this regard, Treagust and Harrison (2000) discussed the aspects of explanations that make up an explanatory framework. They further analyzed Richard Feynman’s Six Easy Pieces in order to identify criteria related to an effective explanation. The researchers categorized explanations into three categories: scientific content explanations, effective pedagogical content explanations, and everyday explanations. Based on philosophical models of explanations, the researchers listed six types of scientific content explanations: Deductive-nomological, deductive statistical, inductive-statistical, complete or comprehensive, causal, and empirical explanations. In addition, they listed six types of pedagogical content explanations: human action, anthropomorphism, teleology, analogy, metaphor, and vignettes. The characteristics of the types of explanations in these categories were discussed throughout the paper with special attention given to pedagogical content explanations. Thus, the researchers were interested in the factors that affect explanation construction, such as, content, context, students, and teachers. After examining Richard Feynman's lecture, 'Atoms in motion', Treagust and Harrison concluded that "effective explanations address in a balanced way the science content, the educational context, teacher factors, and student factors. These factors and available explanatory processes (e.g. deductive, inductive, analogical, etc.) dynamically interact to produce the final explanation" (p. 1167).

Focused on mechanistic reasoning, Russ, Scherr, Hammer, and Mikeska (2008) developed a framework based on the philosophy of science. Russ et al.'s work was meaningful to the practice of explanation construction. In their framework, the researchers focused mainly on one type of explanation – causal mechanisms of natural phenomena and aimed to analyze students' mechanistic reasoning from a philosophical perspective.

Summary

The issue of the lack of a clear conceptualization and articulation of scientific explanation has been a subject of research in science education for over four decades. Researchers in this line of research have attempted to develop various ways, frameworks and theories in order to provide more insight and guidelines to the process of explanation construction. While to this day there is no consensus on a unified definition of a scientific explanation in the philosophy of science, research in science education is still able to benefit from the philosophy of science literature. In many studies, constructing practical guidelines for explanation construction was supported by philosophical and theoretical background. However, these guidelines had some gaps. Thus, there is a need for a framework that takes into account the important theoretical and philosophical background tackled in this line of research. Such a framework should also address the problems that the proposed tools, frameworks and models of examining scientific explanations these studies face.

CHAPTER 3: METHOD

This study was exploratory and qualitative in nature. It aimed to utilize a new explanation-specific framework, which was guided by philosophical models of explanation, to assess freshmen college students', secondary science teachers', and practicing scientists' scientific explanations. In addition, the study sought to characterize the meanings that participants attributed to explanation in the context of science. Two in-depth, semi-structured interviews served as the main source for data collection. In the first interview participants constructed explanations of four different scientific phenomena. In the second interview they provided feedback about the explanations constructed by themselves and other participants. In this study, an in-depth examination and assessment of two of the four scientific phenomena is presented.

Conceptual Framework of the Study: NOSE Framework

Stemming from past science education research on the articulation and conceptualization of explanation in the science classroom, and guided by the philosophical models of explanations, the following section presents the NOSE framework for examining and assessing scientific explanation. NOSE framework builds on previous work on explanation in science education and employs philosophical models of explanation to examine students' scientific explanations. It aims at enabling science education researchers to analyze students' explanations using a framework specific to explanations. At the first developmental stages, NOSE is targeted mainly for science education researchers. More precisely, NOSE framework seeks to enable science education researchers to identify the type(s) of explanation, examine the nature and quality of explanations by integrating various philosophical models of explanations. NOSE framework is sought to help science education researchers to gain a better understanding of the nature of students' scientific explanations and provide a philosophically-grounded approach to examine and assess whether student constructed explanations can be considered explanatory or not. Additionally, for NOSE, both the substance and syntax of the explanation are important.

The following section starts with a general structural definition of what constitutes a scientific explanation. In spite of multiple definitions and rival models of explanations there is a general agreement regarding the idea that an explanation consists of two parts: the first part includes statements about *what* or *how* something is, and the second part includes statements about *why* something is. The first part (the *what*-part) is often not explicitly included in an explanation. Hence, the focus of NOSE framework is on the *why*-part (i.e., the explanatory part) of a scientific explanation. Nonetheless, when present, both parts make up an explanation. This distinction also helps in the demarcation of what is, and is not, explanatory. In other words, and as will be elaborated in this chapter, an answer to a *why*-question that includes only a description of the phenomenon (i.e., the *what*-part) is not considered to be an explanation, in accordance with the NOSE framework.

Scientific Explanation: Structural Definition

A scientific explanation consists of at least two parts: Part 1 can be in the form of a scientific description or a scientific prediction. Part 1 includes statement(s) about *what* something is; whereas part 2 includes statement(s) about *why* something is. Part 1 of an explanation is referred to as the '*What*-Part', and part 2 of an explanation is the '*Why*-Part'. For a statement, or a group of statements, to be considered an explanation, the *Why*-part must be present. On the other hand, the *What*-part of an explanation can be implicit. When explicitly present, the first stage of examining a scientific explanation, NOSE framework suggests identifying the accuracy and relevance of the *What*-part of a scientific explanation. In general, the *What*-Part includes descriptions of what happened or predictions of what will/might happen to the phenomenon-to-be-explained – both of which are discussed in what follows.

The What-Part: Scientific description. In order to assess students' descriptions, the focus is on how students construct their descriptive statements, what features they include in these statements, what they consider important, and whether they represent observable phenomena in a meaningful way. Descriptions of phenomena are not explanatory; instead they include observation statements that are domain-specific. In fact, constructing descriptions in science is closely related to the discipline and its theory (Ford, 2005). As seen in Everback and Crwoley's observation framework, expert scientific observation – that of scientists'- requires more than just sensory observation of phenomena. It provides a link between the observed phenomenon and existing theories (e.g., Everbach & Crowley's, 2009). As defined by Lederman, Abd-El-Khalick, Bell, and Schwartz (2002, p. 500):

> Observations are descriptive statements about natural phenomena that are directly accessible to the senses (or extensions of the senses) and about which observers can reach consensus with relative ease. For example, objects released above ground level tend to fall to the ground.

Thus, observation statements are assessed by their accuracy and relevance to the phenomenon-to-be-explained. For instance, when asking "why did the ball fall to the ground" in the example mentioned earlier, an observation about the color of the ball is considered to be an irrelevant observation. On the other hand, an observation such as "the ball falls to the ground at a constant speed" is an inaccurate observation.

It is important to mention that, in essence, accuracy and relevance are context-dependent criteria. In other words, what counts as relevant varies in content and structure between a 4th grader and a high school student. Hence, the judgement made about the accuracy and/or relevance of observations depends on students' levels, their prior knowledge, and other subjective/qualitative criteria that depend on teacher judgement, context of curricular development, etc. while still maintaining the norms of canonical science. Hence, when examining students' observation statements in science, science educators and teachers should consider these factors and make the necessary judgment.

The What-Part: Scientific prediction. Another type of scientific practice that requests answers to *what* rather than *why* questions includes scientific predictions. Predictions are regarded as statements that posit the consequences of a phenomenon prior to its occurrence. NOSE calls for an examination of whether these statements refer to prior knowledge, and whether or not they include relevant consequences of the phenomenon. Hence, accuracy and relevance are the two key criteria by which students' predictions are assessed. Similar to the case of observations, the accuracy and relevance of a predictive statement depends on the context of teaching and learning.

When examining students' scientific explanations, it is important to distinguish between the descriptive/predictive part(s) (the *What*-Part) and the explanatory part(s) (the *Why*-Part). According to NOSE, The *Why*-Part determines the type, nature and quality of the explanation. Guided by the philosophical work on scientific explanation, NOSE suggests that a scientific explanation includes multiple structural elements that determine its type. Table 3.2 includes a list of the proposed structural elements and their definitions in accordance to the NOSE framework. The nature of these elements and the nature of the interconnection between them determine the *adequacy* of a given explanation. It is important to note that in analyzing students' explanations, researchers need to take into account explainers' levels, their prior knowledge and other subjective/qualitative criteria that depend on teacher judgment, context of curricular development, etc. When applicable, deeper analysis of an explanation or a set of explanations of a phenomenon can be conducted in terms of the explanations' unifying power. Before discussing the structural elements that make up different kinds of explanations, a brief summary of the NOSE types of scientific explanations of natural phenomena is presented.

The *Why*-Part: Types of Scientific Explanations

The nature of the *Why*-Part determines whether an explanation is any one or a combination of the following four types: Deductive nomological (DN), Inductive statistical (IS), Causal and/or Causal Mechanical (CM) explanation. A DN model is one in which a phenomenon is explained by referring it to deterministic/general law(s) and necessary conditions; an IS model is one in which a phenomenon is explained by referring it to probabilistic/statistical law(s); a causal model is one that satisfies Hume's conditions of causality; and a CM model is one in which the conditions of causality are satisfied and necessary causal connections that lead up to the event-to-be-explained are included. Following Weber et al.'s (2013) pragmatic approach, the *Why*-Part of an explanation can include one or more of these models: it can include a deductive model only (hence the explanation is a DN explanation) or a combination of models (e.g., an IS explanation, a CDN explanation, CMDN explanation). More importantly, the type of explanation in the *Why*-Part of an explanation is determined by the nature of the phenomenon at hand. While some phenomena can be subsumed under natural laws, hence be explained by DN explanations; other phenomena are causal and hence require causal explanations. The following sections provides a discussion of how the four major traditional models of explanation can be employed in examining and assessing students' scientific explanations.

Deductive-Nomological (DN) explanation. A DN explanation includes a deductive composition of statements regarding natural phenomena that are logical consequences of general laws of nature. Following Hempel's (1962) work on the covering law model and incorporating solutions to the criticism his model faced, the NOSE framework asserts that a DN explanation should satisfy the following conditions:

1. It should include statement(s) of general laws and natural regularities.
2. It should be a statement of logical deductive form.
3. All and only relevant necessary conditions and all and only relevant information should be included.
4. Irrelevant information should not be included (a solution to the problem of irrelevant premises).
5. In case of causality, only causal derivations are explanatory; derivations from effects are not (a solution to the asymmetry problem).

The DN model has been found to be explanatory for a number of natural phenomena. For example, explanations of the apparent bending of a spoon handle in a glass of water, explanations of the formations of mirror images, explanations of the formations of rainbows, and explanations of the falling of an apple from a tree can be regarded as having a DN character. However, as previously discussed, the DN model has been found to be insufficient at times to fully explain some phenomena. Thus, other models of explanations have been developed for that purpose. Some phenomena are governed by statistical laws and/or causal explanations. More precisely, explanations of phenomena that include terms such as "it is most likely that" or "it will probably be this" are not explained by the DN model.

Inductive-Statistical (IS) explanation. Statements based on laws or theoretical principles of statistic-probabilistic form, or statistical form for short, play a vital role in empirical science (e.g., Hempel, 1965). Hempel's discussion of the Inductive-Statistical (IS) model focuses on how scientific explanations are employed with statistical laws. One important feature about statements of statistical laws is that they are not statements that make claims about only a finite number of cases. Hempel asserted that "law-like sentences, whether true or false, are not just conveniently telescoped summaries of finite sets of data concerning particular instances" (p. 377). An example of a valid statistical law, according to Hempel (1965), is the probabilistic law of radioactive decay. This law is not equivalent to a descriptive report of the frequencies a certain event took place in a certain number of observed instances. On the contrary, the law of radioactive decay provides probabilistic links between hypothetically infinite number of cases.

Similarly, in K-12 science curricula, one can find many science topics of statistical nature. The kinetic molecular theory of gases gave rise to classical statistical mechanics; Brownian motion involves probabilities that are both theoretically and practically definitely less than one; Mendelian genetics provide explanations that are basically statistical. The most dramatic example of statistics in science is the statistical interpretation of the equations of quantum mechanics provided by Max Born and Wolfgang Pauli in 1926-1927; with the aid of this interpretation, quantum theory explains an impressive range of physical facts.

For an IS explanation to be valid, Hempel (1962) requires a high probability (High Probability Requirement or HPR) of a statistical law relative to the event-to-be-explained. According to this requirement, a valid IS explanation must obey HPR requiring a high probability – or a probability close to 1 (more details on the IS explanation is found in Part I of this chapter). However, many philosophers have argued that this is neither necessary nor sufficient for valid statistical explanations (e.g., Salmon, 1998). The more important consideration when examining or constructing an IS explanation is to identify only the factors that are statistically *relevant*. So, if there exists an outcome that is highly probable but unnecessary, then in some of these cases the improbable is more likely to occur. Thus, NOSE framework asserts that an IS explanation should satisfy the following conditions:

1. At least one of the premises involved in the statement must be a statistical or a probabilistic law.
2. The explanation should be of logical inductive form.
3. The explanation must obey the requirement of maximal specificity (RMS). RMS requires all statistically relevant information be included within an explanation (solving the problem of irrelevant premises).

Causal explanation. Similar to Hempel's (1962) view of explanation, Salmon (1998) believed that scientific explanations are indeed answers to *why*-questions. However, Salmon added that not all why-seeking questions are requests for scientific explanation. In particular, causal explanations are answers to why do/does rather than why should. NOSE framework suggests assessing causal explanations using David Hume's (1985) features in causal situations: (1) the temporal precedence of the cause to the effect; (2) the spatiotemporal proximity of the cause to the effect; and (3) constant conjunction – the condition that every time the cause occurs, the effect follows. Thus, NOSE framework asserts that an adequate causal explanation should satisfy the following conditions:

1. Temporal precedence: the cause must always come before the effect
2. Spatiotemporal proximity: the cause and effect must be close to each other (in space and time)

3. Constant conjunction: there must be the same cause-effect sequence on practically *all* observations.

David Hume's three conditions of causality are useful when assessing students' causal explanations. However, there are cases when they are not sufficient. Many philosophers argue that Hume missed an important condition for causal explanations. More specifically, Salmon (1998) argued that Hume was "unable to find any 'necessary connection' relating causes to effects. Such a connection or a series of connections suggests the fourth type of explanation - the Causal-Mechanical explanation.

Causal-Mechanical (CM) explanation. Glennan (2002) affirms that "a mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations" (p. 344). According to Glennan (2002), CM explanations are statements of mechanisms that include traditional accounts of explanations in addition to mechanistic systems. A CM explanation explicates the necessary causal processes and causal interactions that lead up to the event-to-be-explained. Thus, NOSE framework asserts that a CM explanation should satisfy the following conditions:

1. It must adhere to Hume's conditions of causality (temporal precedence, spatiotemporal proximity, and constant conjunction).
2. It must explicate all necessary causal connections (in the form of processes and interactions) that lead up to the event-to-be-explained.

As mentioned in this chapter and stemming from Weber et al.'s (2013) pragmatic approach to explanation, each traditional model of explanations (DN, IS, Causal and CM models) does not seem to be successful, alone, in explaining all types of natural phenomena.

While some phenomena are explained according to a general law and necessary conditions (i.e., DN model), others are strictly causal or causal mechanical. However, some phenomena require the combination of one or more traditional models in order to provide an adequate explanation to the event-to-be-explained. Table 3.1 presents a description of the various types of scientific explanations of natural phenomena. It explicates the general criteria of each type, regardless of their quality. A discussion of the quality of explanations is elaborated later in this chapter.

Structural Elements of a Scientific Explanation

After having synthesized the different types of explanations based on the philosophical models, this section examines the structural elements that make up scientific explanations. Such an examination is important for assessing the quality of scientific explanations constructed by learners. In addition to the analysis of philosophical models of explanation, data collected from the present study helped identify various key elements present in different explanation statements.

NOSE structural elements of an explanation were derived based on the criteria of the types of explanations according to the NOSE framework. Table 3.2 summarizes the different structural elements that make up different explanations. In this table, each structural element is defined, and a corresponding example is presented. For example, lawlike statements, pieces of knowledge, and necessary conditions are structural elements employed in a Deductive-Nomological (D-N) explanation. When identified, a structural element does not indicate whether or not it is accurate, logical, or complete. The identification of a lawlike statement, for example, pertains to the nature rather than the quality (i.e., completeness and adequacy) of the element itself. In general, the mere existence of certain structural elements and the lack of other elements help identify the types of explanations (DN, IS, CDN, etc.). On the other hand, when an in-depth analysis is performed, the quality of these structural elements helps identify the quality of the explanation itself (e.g. Adequate DN, Partially Adequate CDN, in adequate IS, and so on). More structural elements were later added into the process of constructing an explanation map based on participants' constructed explanations that were not intrinsic to the NOSE framework, but were still aligned with the philosophical literature.

Table 3.1

Description of NOSE Framework Types of Scientific Explanation

Type	Description
DN	A DN explanation is one that (1) includes statement(s) of general laws or natural regularities, and (2) is of deductive form.
IS	An IS explanation is one that (1) includes at least one statistical or probabilistic law, and (2) is of inductive form.
Causal	A causal explanation is one that explains the phenomenon (1) as an effect of a cause, (2) within a spatiotemporal proximity, and (3) in which the effect follows the cause.
CM	A CM explanation is one that explains the phenomenon (1) as an effect of a cause, (2) within a spatiotemporal proximity, (3) in which the effect follows the cause, and (4) includes causal connections that lead up to the phenomenon-to-be-explained.
CDN	A CDN explanation is one that explains the phenomenon (1) as an effect of a cause, (2) within a spatiotemporal proximity, (3) in which the effect follows the cause, (4) includes statement(s) of general laws or natural regularities, and (5) is of deductive form.
CIS	A CIS explanation is one that explains the phenomenon (1) as an effect of a cause, (2) within a spatiotemporal proximity, (3) in which the effect follows the cause, (4) includes at least one statistical or probabilistic law, and (5) is of inductive form.
CMDN	A CMDN explanation is one that explains the phenomenon (1) as an effect of a cause, (2) within a spatiotemporal proximity, (3) in which the effect follows the cause, (4) includes causal connections that lead up to the phenomenon-to-be-explained, and (5) is subsumed under natural regularities or general laws.
CMIS	A CMIS explanation is one that explains the phenomenon (1) as an effect of a cause, (2) within a spatiotemporal proximity, (3) in which the cause is most likely (or most probably) followed by the effect, (4) includes causal connections that lead up to the phenomenon-to-be-explained, and (5) is supported by probabilistic or statistical laws.

More specifically, teleological and anthropomorphic statements employed in an explanation were introduced later as structural elements in a given explanation, because a considerable number of participants provided such statements while generating explanations. It is worth noting that at this stage, these structural elements are not set in stone, but rather emergent and responsive to future empirical data that align with philosophical models of explanations.

The Quality of a Scientific Explanation

As has been discussed earlier in this chapter, the nature and type of a scientific explanation are determined by the nature of the structural elements that make up the explanation (e.g., general laws, statistical laws, necessary conditions, causal statements); in addition to the nature of the interconnection between these elements (e.g., simple causal links, deductive reasoning, inductive reasoning). The quality of a scientific explanation, on the other hand, is determined by the quality of the structural elements present in an explanation (accurate relevant pieces of knowledge, irrelevant observations, logical lawlike statements, etc.); in addition to the quality of the interconnection) among these elements (e.g., accurate relevant causal links that satisfy Hume's condition of causality, logical connections).

Moreover, the quality of a scientific explanation, following the NOSE framework, ranges on a continuum of adequate, mostly adequate, partially adequate, and inadequate scientific explanations. Adequacy considers whether or not an explanation is complete, i.e., whether or not it accounts of all relevant components of the phenomenon. Adequacy also includes accuracy and relevance of the structural elements that make up in a given explanation. Hence, for each type of explanation explicated in Table 3.1, criteria for adequacy and completeness are specified. These criteria are pragmatic in nature; i.e., they are dependent on the context in which the explanation takes place, the explainer's prior knowledge, the audience receiving the explanation, and other factors to be determined. The pragmatic nature is of extreme importance as it determines whether or not the given explanation is complete (i.e., depth of explanation), in addition to the relevance and accuracy of the structural elements that make up this explanation. For example, when asked: "When cooking why does a metal spoon get hot than if you use a wooden spoon?", while an explanation given by a 4th grader would be considered complete, that same explanation is incomplete according to a high school student, for example. Thus, the criteria of adequacy explicated in this framework are meant to be applicable to various science teaching and learning contexts. The pragmatic details (i.e., student level, student prior knowledge, purpose of explanation, etc.) are to be determined based on the context of learning.

Table 3.3 through Table 3.10 present descriptions of the criteria of the quality of each type of scientific explanation in accordance with the NOSE framework. For each type of explanation, criteria of completeness, adequacy, relevance, and accuracy of structural elements and their interconnection are presented. For each level of adequacy, an exemplar is also given. Some of these exemplars are hypothetical and generated for the purpose of the framework; while others are examples from this study that were modified to match the quality and type of the explanation at hand.

Explanation Maps

The development of the NOSE framework with its many facets called for the construction of what is termed, an explanation map, for each explanation produced. An explanation map is a visual representation that can help researchers see the "big picture" as they organize statements in an explanation into meaningful connections and visually represent the structural elements that identify the type of explanation generated. Hence, an explanation map includes the various structural elements discussed in this chapter. The construction of explanation maps was adopted from Horn's (1998) and Van Gelder's (2002) definitions of argument maps. Additionally, the process by which explanation maps were constructed followed Martin and Rose's (2008) method of mapping and genre relations. The procedures for constructing an explanation map following the NOSE framework is found in later in this chapter. In addition, Chapter 3 discusses, in detail, how explanation maps were constructed for the purpose of this study.

Methods of Inquiry

Denzin and Lincoln (2005) described qualitative research as a field of inquiry that "crosscuts disciplines, fields, and subject matters" (p. 2). They also emphasized that qualitative research is an activity that "locates the observer in the world" (p. 3). A qualitative approach was suitable for the current study in that it attended to the meanings that participants ascribed to scientific explanations from their stance as individuals, as well as members of groups who approached science from the perspective of science learners, teachers, and practitioners. More specifically, the contexts from which freshmen college students, secondary science teachers, and practicing scientists approached constructing explanations was significantly different in terms of goals, motivations, knowledge, abilities, and skills that these participants brought. Similar to the approach used by Abi-El-Mona and Abd-El-Khalick (2011), a qualitative approach was used in this study because it allowed for understanding of participants' views of explanations by interpreting their answers while taking into account their contextual experiences with science.

Table 3.2

NOSE Proposed Structural Elements that Make Up a Scientific Explanation

Structural Element	Definition	Example
Observation	Observations are descriptive statements about natural phenomena that are directly accessible to the senses (or extensions of the senses) and about which observers can reach consensus with relative ease.	When the lit candle was covered by a glass jar, the water level rose inside the jar.
Inference	Inferences are interpretation based on observations. An inference is not directly available to the senses.	When the lit candle was covered by a jar, water was *pulled* into the jar through the opening.
Prediction	Predictions are regarded as statements that posit the consequences of a phenomenon prior to its occurrence.	After some time, the raisins will all sink to the bottom of the glass and the soda will be flat.
General Law-like statement	General Laws are descriptive statements of relationships among observable phenomena.	This also follows the ideal gas law $PV = nRT$
Probabilistic Law-like statement	Probabilistic laws are probabilistic or statistical descriptive statements of relationships among observable phenomena.	Fick’s law describes the probabilistic behavior of molecules at higher temperatures. And the higher the temperature the higher the probability you are going to have motion.
Piece of Knowledge (PK)	Pieces of knowledge refer to previously learned information, or prior bits of knowledge.	Heat is a form of energy.
Necessary Condition (NC)	A necessary condition is a condition that must be present for an event to occur.	The rugosity of the surface of the raisins allows for entrapment of the bubbles.
Teleological or Anthropomorphic Statements	Anthropomorphic statements include statements that ascribe human feelings and behaviors to elements of the phenomenon-to-be-explained; while teleological statements includes statements of something happening as a function of its end, purpose, or goal.	The water molecules want to stay together. The oxygen part of the water comes into the inside of the jar to feed the flame.
Causal links	A causal link is a statement that explicitly indicates that one event is the result of the occurrence of the other event.	That creates what's called a vacuum, which causes a difference in pressure between the outside and the inside.
Explanatory Connection or Big Idea	A group of structural elements that together form a big idea related to the event-to-be-explained. The elements within a big idea may vary.	The raisins are solid, and they are more compact, and they are heavier than the soda so they sink.
Examples	A comparison between the phenomenon-to-be-explained and another everyday event that according to the explainer highlights respects in which the two are thought to be similar or different	The bubbles act like a floatie in a swimming pool.

Table 3.3

Description of the Quality of a Deductive-Nomological (DN) Explanation

DN Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	- The phenomenon-to-be explained is *not* a logical consequence of the general law or natural regularity included. **OR** - The phenomenon-to-be explained is a logical consequence of the general law or natural regularity included, **BUT** - Only (or mostly) irrelevant and/or inaccurate information (observations, inferences, pieces of knowledge, necessary conditions, etc.) are included.	- Explanation accounts for some but not all components of the phenomenon. - The phenomenon-to-be explained is a logical consequence of the general law or natural regularity included, **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant information (observations, inferences, pieces of knowledge, necessary conditions, etc.) of the phenomenon-to-be-explained.	- Explanation accounts for almost all relevant components of the phenomenon. - The phenomenon-to-be explained is a logical consequence of the general law or natural regularity included. **AND** - Explanation provides mostly statements (observations, inferences, pieces of knowledge, necessary conditions, etc.) that are relevant and accurate with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained.	- Explanation accounts for all relevant components of the phenomenon. - The phenomenon-to-be explained is a logical consequence of the general law or natural regularity included. **AND** - Explanation provides statements (observations, inferences, pieces of knowledge, necessary conditions, etc.) that are relevant and accurate to the phenomenon-to-be-explained.
Exemplar	Water is sucked in because of pressure and where gases are. Gases that are less dense go to the top. And oxygen is less dense than other gases in the air, so it goes to the top of the jar.	After the candle went out all the water kept rising because the pressure inside decreased. The water was under atmospheric pressure, but as soon as the candle started burning it decreased the pressure in some way. In nature there is tendency for things to go from high pressure to low pressure. And the water outside contains some oxygen dissolved in it, and when it went inside the oxygen that was inside the water was used up in some way, and it is trying to come out of water in the form of bubbles.	For combustion reaction we have the wax molecular formula which is $C_{31}H_{64}$ and O_2 you created CO_2 and water. After balancing the equation, you find that you use 47 moles of oxygen for every one mole of the candle wax. So inside you have the CO_2 and the water, and this water is taking up the volume of the O_2 that was used. But you do not have a conservation of volume. The moles of gases of each is not the same. And the fact that you have heated the space you are not dealing with STP.	Light travels at different speeds in different mediums, so in air and in water it is going to travel at different speeds. When that happens you get this effect where light rays will bend. So, if you have multiple rays of light hitting the water in a certain direction we can consider a single point on the water where all the beams of light hit the water at which light is not going to be traveling as quickly as other points. So instead of showing up here it is going to be lagging behind a bit. And this is true also for a point on any line hitting the water.

Table 3.4

Description of the Quality of an Inductive-Statistical (IS) Explanation

IS Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	- The phenomenon-to-be explained does not support the statistical/probabilistic law included. **OR** - The phenomenon-to-be explained supports the statistical/probabilistic law included, **BUT** - Only (or mostly) irrelevant and/or inaccurate information (observations, inferences, pieces of knowledge, necessary conditions, etc.) are included.	- Explanation accounts for some but not all components of the phenomenon. - The phenomenon-to-be explained supports the statistical/probabilistic law included, **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant information (observations, inferences, pieces of knowledge, necessary conditions, etc.) of the phenomenon-to-be-explained.	- Explanation accounts for almost all relevant components of the phenomenon. - The phenomenon-to-be explained supports the statistical/probabilistic law included, **AND** - Explanation provides mostly statements (observations, inferences, pieces of knowledge, necessary conditions, etc.) that are relevant and accurate with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained.	- Explanation accounts for all relevant components of the phenomenon. - The phenomenon-to-be explained supports the statistical/probabilistic law included, **AND** - Explanation provides statements (observations, inferences, pieces of knowledge, necessary conditions, etc.) that are relevant and accurate to the phenomenon-to-be-explained.
Exemplar	When released from above the ground, a ball is most likely to fall to the ground because it is round in shape and because its mass is 1kg.	The raisins sink in the soda simply because it is now favorable to sink in the soda because of their density.	This is a probability related thing. This does not have to happen since the molecules are bouncing around randomly. You could have a situation where all the molecules on the inside just randomly bounce downwards and push all the water out, but that is extremely unlikely.	The two liquids tend to mix together in terms of sheer probability and what we call Boltzmann statistics. The most statistically probable is that the molecules of food coloring are spread out evenly more or less. Temperature is simply how much things move in a given material. And hot water means that your molecules are vibrating or moving around quicker, then it takes less time in the hot water for that food coloring to spread. Fick's law describes this behavior. And the higher the temperature the higher the probability you are going to have motion.

Table 3.5

Description of the Quality of a Causal Explanation

C Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	- The components of the phenomenon-to-be-explained are not logical results of the cause-effect relationship(s) included; i.e., causal links included do not adhere to all Hume's conditions of causality: (1) the cause does not necessarily precede the effect, (2) the cause and effect are not close to each other in time and space, and/or (3) the effect does not always follow the cause.	- Explanation accounts for some but not all components of the phenomenon. **AND** -Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant statements of the cause-effect relationships; i.e., some but not all causal links included adhere to all Hume's conditions.	- Explanation accounts for almost all relevant components of the phenomenon. **AND** - Explanation provides mostly cause-effect relationships that are relevant and accurate and that satisfy Hume's conditions of causality with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained and/or do not satisfy all Hume's conditions of causality.	- Explanation accounts for all relevant components of the phenomenon. **AND** - All logical cause-effect relationships that explain the phenomenon are included; i.e., all the causal links included adhere to all of Hume's conditions.
Exemplar	In the first jar that we used I noticed there were bubbles that caused the flame to put out.	When you put the jar over the candle, this causes vacuum.	Once you cover the candle water starts rising and the candle starts going out. Since the jar is over the top then you have a closed system so there is no more air getting in so the candle will combust all the oxygen that is available.	A force of certain magnitude exerted on an object of a certain mass, caused the object to move a certain distance.

Table 3.6

Description of the Quality of a Causal-Mechanical (CM) Explanation

CM Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	- The components of the phenomenon-to-be-explained are not logical results of the cause-effect relationship(s) included; i.e., causal links included do not adhere to all Hume's conditions of causality. **AND** -Explanation includes irrelevant/inaccurate causal connections that lead up to the phenomenon.	- Explanation accounts for some but not all components of the phenomenon. -Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant statements of the cause-effect relationships; i.e., some but not all causal links included adhere to all Hume's conditions of causality. **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant necessary causal connections that lead up to the phenomenon are included.	- Explanation accounts for almost all relevant components of the phenomenon. - Explanation provides mostly cause-effect relationships that are relevant and accurate and that satisfy Hume's conditions of causality with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained and do not satisfy all Hume's conditions of causality. **AND** - Most necessary causal connections that lead up to the phenomenon are included.	- Explanation accounts for all relevant components of the phenomenon. - All logical cause-effect relationships that explain the phenomenon are included; i.e., all the causal links included adhere to all of Hume's conditions of causality. **AND** - All necessary causal connections that lead up to the phenomenon are included.
Exemplar	When a metallic spoon gets heated, heat transfers in it in the form of heat particles that vibrate and move through the spoon. Each particle hits the one next to it until all of them are heated, and that's why the spoon gets hot all over.	When you put the jar over the candle it creates a vacuum so it brings the water in. Vacuum causes suction.	The raisins first sink to the bottom because of their mass. Then the bubbles adhere to the raisins, causing the raisins with the bubbles to float. At the top, the bubbles pop.	The raisins first sink to the bottom. Then the gas bubbles adhere to the raisins, causing the raisins with the bubbles to float. Gas bubbles have large volume but negligible mass. So the raisins and bubbles together can float to the top. At the top, the bubbles pop so the raisins are just the raisins again and they fall back to the bottom. And the process repeats until there are no more bubbles in the cup.

Table 3.7

Description of the Quality of a Causal Deductive-Nomological (CDN) Explanation

CDN Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	- The phenomenon-to-be explained is not a logical causal consequence of the general law or natural regularity included. **OR** - The phenomenon-to-be explained is a logical causal consequence of the general law or natural regularity included, **BUT** - Only (or mostly) irrelevant and/or inaccurate information (observations, inferences, pieces of knowledge, necessary conditions, causal links, etc.) are included.	- Explanation accounts for some but not all components of the phenomenon. - The phenomenon-to-be explained is a logical causal consequence of the general law or natural regularity included, **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant information (observations, inferences, pieces of knowledge, necessary conditions, causal links, etc.) of the phenomenon-to-be-explained.	- Explanation accounts for almost all relevant components of the phenomenon. - The phenomenon-to-be explained is a logical causal consequence of the general law or natural regularity included, **AND** - Explanation provides mostly statements that are relevant and accurate with only a few statements that are inaccurate and/or irrelevant.	- Explanation accounts for all relevant components of the phenomenon. - The phenomenon-to-be explained is a logical causal consequence of the general law or natural regularity included, **AND** - Explanation provides statements that are relevant and accurate to the phenomenon-to-be-explained.
Exemplar	Water moves on a gradient from high concentration to low concentration. So there was more water outside the jar than inside the jar. Once you put the jar over the candle, you already had a little bit of water trapped inside, there was all the oxygen molecules on top of the water. Once the fire started burning through those oxygen molecules, the atmosphere within the jar, which was in a sense pushing down on the water was removed. And since water on that gradient, it moves from high concentration to low concentration, it started moving up because now it had more space to move up.	In this experiment, the pressure inside with the candle has drastically reduced, because now all the water is being pushed by the atmospheric pressure outside. And the pressure inside is reduced because clearly there was a good amount of oxygen in the jar. The flame burnt it into some sort of soot. The mass is going to be the same. Gases are particles that are moving around very quickly. But things like smoke and soot and heavier. So a lot of stuff that was bouncing around is now turned it into something that is moving a lot slower.	The raisins first sink to the bottom because of their density. And the bubbles cause the raisins to become less dense. Less dense objects float.	A force of certain magnitude exerted on an object of a certain mass, causes the object to move a certain distance with a certain acceleration. This follows Newton's second law of motion where, as the force increases the acceleration also increase if we maintain a constant mass.

Table 3.8

Description of the Quality of an Causal Inductive-Statistical (CIS) Explanation

CIS Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	- The phenomenon-to-be explained does not support the causal statistical/ probabilistic law included. **OR** - The phenomenon-to-be explained supports the statistical/probabilistic law included, **BUT** - Only (or mostly) irrelevant and/or inaccurate information are included.	- Explanation accounts for some but not all components of the phenomenon. - The phenomenon-to-be explained supports the statistical/probabilistic law included, **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant information of the phenomenon-to-be-explained.	- Explanation accounts for almost all relevant components of the phenomenon. - The phenomenon-to-be explained supports the statistical/probabilistic law included, **AND** - Explanation provides mostly statements that are relevant and accurate with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained.	- Explanation accounts for all relevant components of the phenomenon. -The phenomenon-to-be explained supports the statistical/probabilistic law included, **AND** - Explanation provides statements that are relevant and accurate to the phenomenon-to-be-explained.
Exemplar	When released from above the ground, a ball is most likely to fall to the ground because gravity causes it to fall.	The raisins sink in the soda simply because it is now favorable to sink in the soda because of their density. And density causes things to sink or float.	Since the molecules are bouncing around randomly. You could have a situation where all the molecules on the inside just randomly bounce downwards and push all the water out, but that is extremely unlikely. So now that the temperature is increased, this causes the water to be drawn into the jar. This is the most likely situation.	Radioactive decay is caused when an unstable atomic nucleus spontaneously breaks into smaller more stable fragments. Carbon 14 atoms decay in a statistically regular pattern providing a technique for radiocarbon dating. Other radioactive atoms decay with different statistical patterns. One of the implications of these statistical regularities is that there exists a high probability that a given tritium atom, for example, will decay in a period of 5715 years – that is, there is 50% chance that a given carbon 14 atom will decay in the same period, and there is a small probability that a given Uranium 238 atom will decay in that same period.

Table 3.9

Description of the Quality of a Causal-Mechanical Deductive-Nomological (CMDN) Explanation

CMDN Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	-The phenomenon-to-be-explained is subsumed under irrelevant/inaccurate natural regularities or general laws. - The components of the phenomenon-to-be-explained are not logical results of the cause-effect relationship(s) included; i.e., causal links included do not adhere to all Hume's conditions of causality. **AND** -Explanation includes irrelevant/inaccurate causal connections (in the form of processes and interactions) that lead up to the phenomenon.	- Explanation accounts for some but not all components of the phenomenon. - The phenomenon-to-be-explained is subsumed under logical natural regularities or general laws. -Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant statements of the cause-effect relationships included; i.e., some but not all causal links included adhere to all Hume's conditions of causality. **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant necessary causal connections (in the form of processes and interactions) that lead up to the phenomenon are included.	- Explanation accounts for almost all relevant components of the phenomenon. - The phenomenon-to-be-explained is subsumed under logical natural regularities or general laws. - Explanation provides mostly cause-effect relationships that are relevant and accurate and that satisfy Hume's conditions of causality with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained and do not satisfy all Hume's conditions of causality. **AND** - Most necessary causal connections (in the form of processes and interactions) that lead up to the phenomenon are included.	- Explanation accounts for all relevant components of the phenomenon. - The phenomenon-to-be-explained is subsumed under logical natural regularities or general laws. - All logical cause-effect relationships that explain the phenomenon are included; i.e., all the causal links included adhere to all of Hume's conditions of causality: 1) the cause always precedes the effect, (2) the cause and effect are always close to each other in time and space, and (3) there is always the same cause-effect sequence on practically all observations. **AND** - All necessary causal connections (in the form of processes and interactions) that lead up to the phenomenon are included.
Exemplar	The raisins initially sink then they float back up. The carbonation and the molecules within the soda are interacting within the raisins. So there seems to be an exchange at the membrane level, where initially it goes in and this leads to a decrease in density, which makes the raisins rise. The water molecules are able to go into and out of the raisin.	The raisins start off being heavier, but because of the bubbles they are lifting the raisins. And then once the carbonation or the bubbles come off, they go back down to sinking. And that's	At first the raisins fell to the bottom but then they started floating up and down due to the bubbles in the soda. They go up due to the carbonation pushing them up. So they first sink to the bottom due to the force of the dropping of the raisins. It broke the surface tension of the surface and fell to the bottom.	By placing the jar we limit the amount of oxygen available for the candle. And as it burns through the oxygen then the amount of gas is less, but still within the same volume of the jar. But the water level does not rise instantly but continues to rise after the flame goes out.

(Table continues)

Table 3.9 (Continued)

CMDN Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Exemplar (Cont'd)	And you can see bubbles surrounding the raisin. These are water bubbles. And it is also that CO_2 is traveling in and out of the raisins. CO_2 should be able to make it across the membrane of the raisin. When enough bubbles escape the raisins, the raisins go to the bottom. So that tells me that the bubbles are causing the raisins to go up. And as more bubbles surround the raisin, you are seeing principles of cohesion and adhesion forming this hydration shell that is leading to decrease in density of the raisin. So then it floats	because the bubbles are carrying it up and making it lighter so they come up there. And the bubbles come off when they hit the surface, so they are heavier, so they come back down. And the process will keep going on like this.	They also fell because of their weight, that is their weight compared to the amount of the space that they take up. The raisins are just heavy enough to stay at the bottom. And the bubbles are providing some additional buoyancy so they go up. The gas in the fluid is trying to get out of the fluid, and it must be wanting to disperse. The bubbles are carbonation in the fluid, so due to entropy that is expanding and the gas diffusing out they create their own surface tension. And because bubbles are filled with this gas they are lighter than the fluid and when stuck to the raisins they are giving it extra buoyancy that is causing it to go to the top. Once they reach the surface some of those bubbles pop probably because they are hitting that surface tension at the top of the liquid so they are no longer assisting the raisin in floating.	So we have a gas that becomes lighter because there is less of it. And according to the equation $PV = nRT$. So it is not the volume maybe it is temperature since it is $PV= nRT$. So the gas would expand because of the temperature increase. So as T increases V increases in theory. So the gas is being heated up by the flame. And the hotter the gas is the more room it takes. And when the flame disappears there is no source of heat anymore, and the gas is going to cool down. And by cooling down it is going to retract.

Table 3.10

Description of the Quality of a Causal-Mechanical Inductive-Statistical (CMIS) Explanation

CMIS Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
Description	-The phenomenon-to-be-explained supports irrelevant/inaccurate statistical/probabilistic laws. - The phenomenon-to-be explained does not support the causal statistical/ probabilistic laws included; i.e., causal links included do not adhere to all Hume's conditions of causality. **AND/OR** -Explanation includes irrelevant/inaccurate causal connections (in the form of processes and interactions) that lead up to the phenomenon.	- Explanation accounts for some but not all components of the phenomenon. - The phenomenon-to-be-explained supports accurate and relevant statistical/probabilistic laws. -Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant statements of the cause-effect relationships included; i.e., some but not all causal links included adhere to all Hume's conditions of causality. **AND** - Explanation provides mixed accurate and/or inaccurate, and relevant and/or irrelevant necessary causal connections (in the form of processes and interactions) that lead up to the phenomenon.	- Explanation accounts for all relevant components of the phenomenon. - The phenomenon-to-be-explained supports accurate and relevant statistical/probabilistic laws. - Explanation provides mostly cause-effect relationships that are relevant and accurate and that satisfy Hume's conditions of causality with only a few statements that are inaccurate and/or irrelevant to the phenomenon-to-be-explained and do not satisfy all Hume's conditions of causality. **AND** - Most necessary causal connections (in the form of processes and interactions) that lead up to the phenomenon are included.	- Explanation accounts for all relevant components of the phenomenon. - The phenomenon-to-be-explained is subsumed under logical natural regularities or general laws. - All logical cause-effect relationships that explain the phenomenon are included; i.e., all the causal links included adhere to all of Hume's conditions of causality: 1) the cause always precedes the effect, (2) the cause and effect are always close to each other in time and space, and (3) there is always the same cause-effect sequence on practically all observations. **AND** - All necessary causal connections (in the form of processes and interactions) that lead up to the phenomenon are included.
Exemplar	NA	NA	The gas molecules in the jar are bouncing around randomly all the time, and the force exerted when it is bouncing around is determined by the temperature of the gas. So the air outside the jar is going to end up pushing down on the water and it is going to	A glass of ice water melts in air at room temperature because the difference in temperature between the room (i.e., the surroundings) and the cold glass of ice and water starts to equalize as portions of the thermal

(Table continues)

Table 3.10 (Continued)

CMIS Quality	Inadequate	Partially Adequate	Mostly Adequate	Adequate
			push some of the water up into the jar until the forces equal. This comes down to the relevant equation for an ideal gas is PV is proportional to T. So after the flame goes out the temperature will slowly start dropping and when that happens the water level rises more. When the temperature decreases that means that air inside slowly has less and less energy. As it is striking the surface of the water it is striking with less and less force. And so this counter-balances the pressure from the outside and it ends up drawing more water into the jar. This is a probability related thing: this does not have to happen since the molecules are bouncing around randomly. You could have a situation where all the molecules on the inside just randomly bounce downwards and push all the water out, but that is extremely unlikely.	energy from the surrounding (the warmer system) spread to the cooler system of the glass of ice. When time passes, the temperature of the glass and its contents and the temperature of the room will be equal. That is, the entropy of the room has decreased as some of its energy is transferred to the ice and water. This is because the entropy of the system of ice and water, which is a measure of how far the equalization has progressed, has increased more than the entropy of the surrounding room has decreased.

Participants

The study was undertaken with three groups of participants in a large, Midwestern University and neighboring communities: freshmen college students, secondary science teachers, and practicing scientists. A total of thirty participants, ten from each group, were invited to take part in the study. Appendix A presents the letters and informational flyers that were used to invite participation in the study. Thus, participants were self-selected and included those who agreed to voluntarily participate in the data collection activities. Informed consent (see Appendix B) were secured from all participants prior to their involvement with the study.

First, a call for participation was sent through invitation letters to freshmen students at the participant University following due procedures associated with accessing students for research purposes. In addition, informational flyers were posted on social media (Facebook and Twitter) calling for participants, as well as printed and posted on walls and bulletin boards across campus and the surrounding area as approved by the Institutional Review Board (IRB). In order to ensure that all participating freshmen students had a high school science background, the calls for participation indicated that freshmen students who wished to participate in the study should have completed at least two years of high school science. Appendix A includes the sample invitation letters and sample informational flyers used to call for study participants. A total of 15 individuals signed up but three declined to continue with the second interview due to exams and family circumstances, while two were unable to provide explanations to at least three of the four scenarios included in the study. Their answers to questions mainly included statements such as “I don’t know”, “I am not sure”, etc.. Due to the nature of the study, providing explanations was necessary for conducting the second set of semi-structured interviews. Thus, they were not invited to participate in the second interview. Hence, ten freshmen students – six males and four females – were 18 years old and were selected to participate in this study. Four of the 10 students had enrolled in advanced placement and honor high school science courses. Two of those four students had relevant experiences in science outside their schools: one had participated in a national science fair event in high school, while another had participated in a summer outreach program in physics. Two additional students were members of science clubs in high school but had not enrolled in advanced placement or honor high school science courses

Second, ten participant secondary science teachers (5 Males, 5 Females) were accessed through the Office of School University Research Relations. All participant teachers held a BS in chemistry, physics, cellular biology, environmental biology, animal sciences, or natural resource science and taught high school physics, chemistry, biology, physical science, environmental science, science and agriculture, and earth science. Six teachers held an MS degree in education (secondary education, agricultural education, or educational administration), one had earned her MS degree in chemistry, and one in clinical psychology. Their ages ranged from 25 to 60 years (M = 42.8) and their teaching experience ranged from 1 to 39 years (M = 13.7). All 10 teachers noted their participation in relevant experience in science education-related projects outside of school. Three teachers had ongoing participation with campus-based professional development projects that included developing interactive science instructional videos and other projects relating science with engineering. Two teachers had participated in summer camps for middle and high school students, developed science lessons, and assisted in developing science curricula. One teacher had developed a physics course with the Physics Department at the same Midwestern University in which the study took place. The course was aimed to assist students who were interested in developing a conceptual understanding of the world around them. Another teacher had participated in professional development workshops in education and agriculture. One teacher had another part time job and worked at a vet clinic and one teacher was also a forestry technician.

The third participant group consisted of ten practicing scientists. In this study, and similar to Abi-El-Mona and Abd-El-Khalick's (2011) study, practicing scientists were defined as advanced graduate science students in the final stages of their doctoral program (i.e., in the dissertation research and/or writing phase), postdoctoral fellows, and professional scientists. Participants were approached through the Office of the Chancellor and through the scientists' research laboratories. Hence, 10 practicing scientists (four males, 6 females) participated in this study. Their age ranged from 26 to 36 (M = 28.5). All participant scientists were either working or studying at the same University where the study was conducted. Two of the 10 scientists were post-doctoral scholars: one in Bio-Chemistry and another in Nutrition; while the remaining eight were doctoral candidates in the last stages of their dissertations. Doctoral candidates' content experience ranged over a variety of topics including physics, mechano-chemistry, material science and engineering, soft material sciences, condensed matter theory, and neurotoxicology. Practicing scientists' research interests included studying polymer reaction under force, machine learning to analyze particle collision data, quantum physics of crystals at low temperatures, behavior of metals in strong magnetic fields, deformation of metals at a very small scale, how bacteria use oxygen for metabolic processes, examining bacteria that create carcinogenesis, and examining how biological molecules arrange themselves into complicated structures.

Procedures

Phase I. Participants generated explanations for four scientific scenarios. The study was conducted in three phases. In the first phase, a semi-structured individual interview (Interview I) was conducted with all participants. Interview I was comprised of four scenarios targeting explanations related to everyday scientific phenomena. The scenarios varied between predict-observe-explain (POE) type activities and explain-only questions. In the first interview, the POE-type scenarios were based on discrepant events that aimed to elicit participant's curiosity and encourage them to provide scientific explanations to the phenomenon at hand. Participants were first asked to predict and explain the possible phenomenon, then observe the phenomenon taking place and provide an explanation post observation. Participants' explanations took into account their observations, especially in terms of supporting or contradicting their predictions. The latter possibilities, at times, generated the need for additional or alternative explanations. Two popular hands-on POE activities were used: The Dancing Raising Scenario (See Scenario I of Appendix C) and the Burning Candle Scenario (See Scenario II of Appendix C). In addition to the POE-scenarios, participants were asked to explain two phenomena from daily life: a penny in a water bucket demonstration and food coloring in hot and cold-water demonstration (See Scenarios III and IV in Appendix C). After sharing their responses, follow-up questions were used to probe participants' ideas and clarify any ambiguities in their answers. At the end of each scenario, participants were asked to provide a final explanation of the phenomenon at hand. During the first round of interviews, participants were also asked whether or not they were familiar with the phenomena at hand. While it did not constitute a full-blown assessment of their prior knowledge, participants' self-reported datum was factored into the analysis. Hence, each participant generated four final explanations (one per scenario) during the first interview. With a total of 30 participants, with 10 members in each of the three groups, a total of 120 final explanations were generated. The present study focuses on participants' final explanations from the first two scenarios: The Candle in a Jar (CIJ) scenario and the Dancing Raisins (DR) scenario. Hence, a total of 60 explanations were used for the analysis of this study. The full protocol for Interview I appears in Appendix C. All interviews were audiotaped and transcribed verbatim for analysis. A typical interview lasted between 40 and 50 minutes.

Phase II. Researcher generated explanation maps. In this phase, the researcher coded each participant's interview transcripts, and used the coded transcripts to generate a corresponding explanation map of the final explanation provided by the participants following procedures of the NOSE framework. The detailed process of constructing these explanation maps, which took about one month, appear later in this chapter, and the procedures undertaken to ensure the consistency and reliability of the construction are detailed below (See "Phase I" under the Data Analysis subsection). The explanation maps were used during the third phase of the study. The section below presents a summary of explanation maps and how such maps are constructed for the purpose of this study.

Similar to argument mapping developed by Horn (2003), explanation mapping is basically constructing a diagram that represents the explanation, identifying the nature of its structural elements and the nature of the interconnection between these elements. Such a representation facilitates the determination of the type, nature and quality of the corresponding explanation. Hence, an explanation map is a graphical representation of the structural elements involved in an explanation. Similar to flowcharts, explanation maps are shape-and-arrow diagrams. In this study, following Horn's (2003) procedure of argument mapping, explanation mapping was used to represent the directional flow of participants' thoughts (taken from the audiotapes and transcripts). For the purpose of examining participants' explanations, Martin and Rose's (2008) explanations genre relations and common conjunctions were adapted to the construction of explanations maps from participants' transcriptions.

In this study, in Part I of data analysis, after identifying an explanation of a *why*-question, the researcher constructed explanation maps directly from participants' audiotapes and transcripts following the NOSE framework. Actual interview statements were used and placed in various shapes. Each explanation map was represented by a diagram constructed that allowed the explanations to be visible. Inspiration ® software version 9.2.2 was used to build the map due to the flexibility provided by this software. The procedure used is described below.

Procedure for creating an explanation map starting with a participant's interview transcript using nose framework

1. The transcript was closely read to identify the final explanation and the relevant parts of the explanation. At this step a final explanation was the answer to the why-question asked by the researcher during the first round of interviews. All statements associated with the particular answer to the why-question were included within the same explanation.
2. Each explanation map was preceded by a sequence diagram similar to the ones constructed by Martin and Rose (2008). These diagrams were helpful in constructing explanation maps. An example of the sequence diagram appears in Figure 3.1.
3. The first direct answer made by the interviewee was identified (e.g., "When you added raisins to the glass of 7UP the rains sank."). This first direct answer was assigned with "1" in an explanation map, and subsequent statements were given subsequent numbers. Hence, statements in an explanation map were numbered by chronological order following the participant's oral articulation of his/her explanation. In this study, participants provided verbal rather than written explanations, and their statements belonging to the same idea were sometimes expressed at different times during their

explanation articulation. Thus, statements of an explanation that belonged to the same idea were positioned together in an explanation map.

4. Map construction began with the direct answer - set in a box at this stage - made by the participant in regards to the *why*-question at hand.
5. Links were constructed in relation to the direct answer based on the flow of the interviewee's explanation and following Martin and Rose's (2008) explanations genre relations and common conjunctions. These links were demonstrated using arrows with tips indicating the direction of the explanation being generated.
6. Conjunctions linking direct answers were identified and added into the arrows. Exact conjunctions were used when stated explicitly by the participant. In cases where no conjunction was articulated by the explainer, an inferred conjunction was added between parentheses based on the context of the explanation.
7. Statements in boxes constructed in Step #3 were then identified into their corresponding structural elements (e.g., observations, inferences, pieces of knowledge), and they were modified into their corresponding shapes. (A) Statement(s) in a certain shape corresponded to one specific idea. For example, "After some time, the raisins will all sink to the bottom of the glass and the soda will be flat" was identified as a prediction and was included in a circle circumcised in a square (for a complete list of all shapes of structural elements, see Table 3.11).
8. As listed in Table 3.11, causal links were bolded arrows that specified an explicit cause-effect relationship between two or more structural elements in an explanation map. Causal links were identified by the conjunctions and other linguistic identifiers used by the explainer. In fact, Martin and Rose's (2008) list was used to identify the type of

connections between (among) any two (or more) structural elements. The list included common conjunctions that help identify different types of explanations. For example, under the right contextual circumstance, words such as *cause*, *because, so, therefore*, etc. helped identify causal relations and/or deductive explanations. Conjunctions such as *although, even though, but, however,* etc. indicated unexpected consequences of relations. Conditional explanations (which are similar to Inductive-Statistical Explanation in the NOSE framework) included words such as *if, then, provided that, as long as* for expectant conditional explanations.

9. In addition to the conjunctions, criteria that were unique to each type of explanation were identified following procedures of the NOSE framework. This helped in determining the type of explanation at hand. For more details on the criteria of each type of explanation, see Table 3.2. For example, a D-N explanation was identified through law-like statements, general laws, necessary conditions, and related pieces of knowledge. On the other hand, a Causal explanation was identified through simple cause-effect links in expressions such as *because, this leading to, this causes, effect/affect, etc.* Recall that the structural elements among shapes helped identify one or multiple types of explanation.
10. A explanation map that corresponded to the sequence diagram was finally constructed. Figure 3.2 represents the explanation map that corresponds with the coded transcript in Figure 3.1.

Coding transcripts. Each transcript of a final explanation—from Interview I—was first divided into its different statements joined by conjunctions and linking words. These statements were then coded according to the structural elements of the NOSE framework (see Table 3.11). The excerpt below is an example of an explanation provided by Faith (pseudonym), a College

Freshman student, explaining why water rises after covering a lit candle with a jar. Figure 3.1 presents the corresponding sequence diagram that included the coding of the transcript in which the conjunctions and the structural elements in the explanation were identified:

> When you put the water on the plate with the candle, and you lit the candle, and put the jar on top, the water gets sucked into the jar and then the candle will burn out. And the water inside the jar will bubble but it won't leave the jar. It will stay in it. When you put the jar over the candle, it creates a vacuum, so it sucks in the water. The flame goes out because it needs Oxygen to keep burning. And when you put the jar over it, it takes away its source of Oxygen because it will burn all of it that is in the jar.

It is worth noting that the exact conjunctions were used when stated explicitly by the explainer. However, in cases where no conjunction was articulated by the explainer, an inferred conjunction was added between parentheses based on the context of the explanation. In addition, as discussed in Chapter 2, the structural elements of an explanation were initially derived based on the criteria of the types of explanations according to the NOSE framework. A total of 120 coded transcripts of participants' final explanations were generated.

Constructing explanation maps. The construction of explanation maps was adopted from Horn's (1998) and Van Gelder's (2002) definitions of argument maps. Additionally, the process by which explanation maps were constructed, using the NOSE framework, followed Martin and Rose's (2008) method of mapping and genre relations. First, similar to an argument map, an explanation map is a representation of an explanation "in which the inferential structure is made completely explicit, usually by graphical techniques" (Van Gelder, 2002, p. 85). Rather than using boxes and arrows to indicate claims and evidential relationship in an argument (as is the case in Gelder's argument maps), an explanation map uses various shapes (boxes, bubbles, diamonds, etc.) and arrows to indicate the structural elements and type(s) of explanation at hand in accordance with the NOSE Framework.

[Observation] When you put the water on the plate with the candle, and you lit the candle, and put the jar on top,

(then)

[Inference] the water gets sucked into the jar

and then

[Observation] and then the candle will burn out.

and

[Observation] And the water inside the jar will bubble but it won't leave the jar. It will stay in it.

when

[Necessary Condition] When you put the jar over the candle,

(so)

[Inference] it creates a vacuum,

so

[Inference] so it sucks in the water.

(and)

[General Lawlike statement] The flame goes out because it needs Oxygen to keep burning.

and when

[Necessary Condition] And when you put the jar over it,

(so)

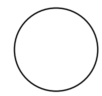
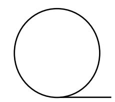
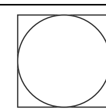
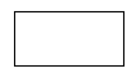
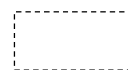
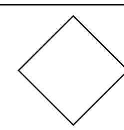
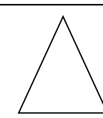
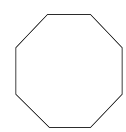
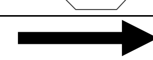
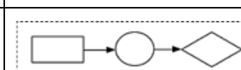
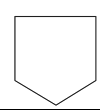
[Inference] it takes away its source of Oxygen because it will burn all of it that is in the jar.

Figure 3.1 Faith's coding of her explanation of the water rising phenomenon. Bolded links in black indicate causality. Conjunctions in parentheses refer to implicit conjunctions inferred.

The coded transcripts were used in the construction of the corresponding explanation maps. Each structural element was denoted by a shape that was arbitrarily chosen—and consistently applied—to help with visually identifying and comparing various types of explanations.

Table 3.11

Shapes and Definitions of Structural Elements in an Explanation Map According to NOSE Framework

Structural Element	Operational Definition	Shape
Observation	Observations are descriptive statements about natural phenomena that are directly accessible to the senses (or extensions of the senses) and about which observers can reach consensus with relative ease.	
Inference	Inferences are interpretation based on observations. An inference is not directly available to the senses.	
Prediction	Predictions are regarded as statements that posit the consequences of a phenomenon prior to its occurrence.	
General Law-like statement	General Laws are descriptive statements of relationships among observable phenomena.	
Probabilistic Law-like statement*	Probabilistic laws are probabilistic or statistical descriptive statements of relationships among observable phenomena.	
Piece of Knowledge (PK)**	Pieces of knowledge refer to previously learned information, or prior bits of knowledge. PK is shorthand for prior knowledge about scientific information.	
Necessary Condition (NC)	A necessary condition is a condition that must be present for an event to occur.	
Teleological or Anthropomorphic Statements	Anthropomorphic statements include statements that ascribe human feelings and behaviors to elements of the phenomenon-to-be-explained; while teleological statements includes statements of something happening as a function of its end, purpose, or goal.	
Causal links	A causal link is a statement that explicitly indicates that one event is the result of the occurrence of the other event.	
Explanatory Connection or Big Idea	A group of structural elements that together form a big idea related to the event-to-be-explained. The elements within a big idea may vary.	
Example	A comparison between the phenomenon-to-be-explained and another everyday event that according to the explainer highlights respects in which the two are thought to be similar or different.	

* Structural element pertaining to probabilistic nature have the same shape as those related to general law elements except that they are dashed (see the dashed rectangular shape of a probabilistic law-like statement in the table.
**The reader is advised not to confuse NOSE pieces of knowledge with diSessa's (1986) knowledge in pieces. The two are very different and serve different purposes.
Note: In explanation maps, elements were marked in bolded black color to indicate that they were irrelevant to the phenomenon at hand. Accuracy was not indicated in the map, since it requires a thorough examination.

Table 3.11 uses the list of structural elements and their definitions from Table 3.2 and adds a key of the shapes of the various structural elements that were consistently used in the analysis. For convenience, definitions of each structural elements are also included in Table 3.11.

An important aspect of explanation maps was that statements in an explanation map were numbered by chronological order following the participant's oral articulation of an explanation. In this study, participants provided verbal rather than written explanations, and their statements belonging to the same idea were sometimes expressed at different times during their explanation articulation. Thus, statements of an explanation that belonged to the same idea were positioned together in an explanation map. For example, all statements associated with the flame of the candle in the CIJ scenario were positioned together in an explanation map even if some of these statements were articulated in the beginning of a participant's explanation generation, while other statements of the same idea were articulated later on during explanation generation. Figure 3.2 represents Faith's explanation map that corresponds with the coded transcript in Figure 3.1.

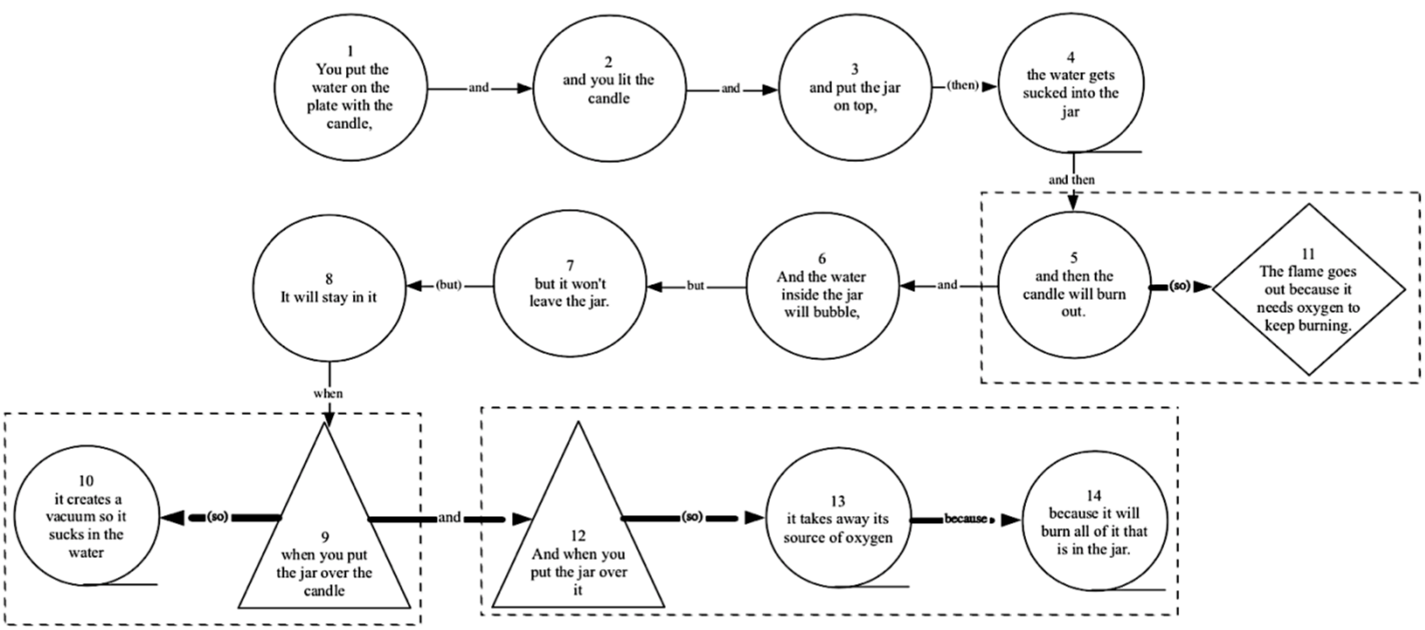


Figure 3.2 Faith's coding of her explanation of the water rising phenomenon. Bolded links in black indicate causality. Conjunctions in parentheses refer to implicit conjunctions inferred.

Phase III. Participants assessed explanations generated by other participants. Interview II was conducted during this phase. In this interview, participants in each group assessed and provided feedback on explanations generated during the first phase by another participant in their own group, as well as one participant from each of the other two groups. Thus, in addition to examining one of his/her own explanation, each participant examined three additional explanations. Hence, each participant assessed and provided feedback on a total of four final explanations and their corresponding explanation maps. In order to ensure a balanced treatment, the assignment of explanations to be assessed was random with no repetition of scenarios. In particular, random assignment ensured that explanations of all scenarios from the first interview were included in the second interview. In addition, due to the relatively small sample size in the study, in the second interview, explanations of the same scenario were not presented to the same participant more than once. For example, in addition to revisiting one explanation (and its corresponding explanation map) from their own explanations, each participant scientist examined a randomly selected explanation of a different scenario from among explanations generated by the other nine scientists, one explanation of a different scenario than the first two selected from those generated by the 10 participant teachers, and one explanation randomly selected from the 10 participant student explanations (also of a different scenario). In this interview, participants were first provided with a transcript of one of the final explanations they provided during Interview I in addition to its corresponding explanation map. Participants were then asked to comment on the accuracy of the map in capturing their explanations. Next participants assessed and provided feedback on final explanations generated during the first phase by one participant from their own group, as well as one form each of the other two groups. The full protocol for Interview II is in Appendix D. Thus, a total of 90 assessments were generated, three by each of the 30 participants.

When examining final explanations generated by other participants for the three groups, group memberships of the explainers were held anonymous in order to avoid any biases that might result from views related to assuming that the explainer had less or more knowledge and expertise compared to the individual assessing the explanation. To do this, the researcher made sure that all phrases that could give clues about the possible background of an interviewee were not disclosed to the participants. For example, transcribed segments that hinted to the identity, years of experience, or type of work of the interviewee were not read by participants assessing the explanations. In addition, this approach helped shield participants' gender identities and ages – additional attributes which could cause bias that could be inferred by an interviewee from listening to an audio recording instead of reading a transcribed segment. These transcribed segments were edited prior to Interview II with each participant.

During Interview II, participants were asked to define scientific explanation in their own words, to assess the quality of the explanations they examined, and justify their assessment. Eventually, the interviewee was asked to judge whether an explanations was 'valid' or 'adequate', and 'complete' or 'incomplete.' Interviewees were also asked to choose what they considered to be the 'best' explanation from among the four explanations they examined (including their own generated explanation) regardless of the phenomenon that was explained. Finally, participants were asked to list, in their own words, the criteria they used to assess the validity or adequacy of these explanations. In addition, the interviewer asked probing questions that aimed to elucidate the interviewee's implicit criteria used to assess or judge the explanations at hand. All interviews, which lasted about 45 minutes, were audiotaped and transcribed for analysis. The interview protocols detailed in Appendices C and D guided Interviews I and II respectively. However, unplanned follow-up and probing questions were also used during the interviews.

Pilot Study

The aforementioned procedures were tested in a pilot study, which took place about one semester prior to data collection. For convenience, the pilot study involved a sample of five participants: two undergraduate students (a freshman student and a junior student), one community college science instructor, and two practicing scientists. In accordance with the above procedures, each of the pilot study participants was interviewed twice. At the conclusion of each interview, participants were asked to comment on the interview procedures as a whole, and on the clarity of the tasks they were assigned, as well as the questions asked of interviewees about the explanations they examined. Participants' responses in the pilot study were used to rephrase and improve any unclear questions in the interview protocols.

Analysis of Data

Data were analyzed in three phases: The first phase involved analyzing participants' explanation maps following the NOSE framework. The second phase involved analyzing transcripts generated during the second interview to characterize participants' perceptions of the nature of explanations and derive the criteria they developed to judge the 'quality' or 'goodness' of explanations. This latter analysis was followed by comparing and contrasting the analysis within and across the three participant groups. The third phase of analysis focused on comparing the criteria derived from the second phase with aspects of explanations emphasized in NOSE framework. Details on the procedures that were followed for data analysis are presented in the following sections.

Part I: Analyzing Participants Final Explanations. Part I of data analysis involved identifying participants' final explanations from the first interview transcripts, generating their corresponding explanation maps, and analyzing them using the NOSE framework. Recall that at the end of each scenario during the first interview, participants were asked to provide a final explanation of the phenomenon at hand. Before moving onto the next scenario, the interviewer asked participants for a final wrap up in which they described what happened in the activity and explained why it happened. These final explanations were used in the second interview.

So, this part of data analysis involved the construction of explanation maps prior to conducting the second round of interviews. Maps of final explanations were constructed from participants' transcripts generated during the first round of interviews. Thus, following the NOSE framework, all final explanations were identified, and participants' transcribed explanations were used to produce explanation maps of final explanations.

Establishing inter-coder reliability. In order to ensure consistency and reliability of explanation map analysis in a way in which the NOSE framework accurately analyzed the transcribed explanation, a post-doctoral scholar in science education teamed with the researcher to analyze 10% of the explanation maps constructed by the researcher. The scholar held a BS degree in chemistry and had taught science at the pre-college and college levels. The scholar had no prior exposure to, and was unfamiliar with, the NOSE framework. The two researchers met several times in order to familiarize him with the framework. The researcher then provided the post-doctoral scholar with 12 explanation maps to analyze according to the NOSE framework. For this purpose, explanation maps of the four scenarios were randomly selected from the three participating groups: one map was randomly selected form each of the three participating groups for each of the four scenarios. For example, one explanation map of the Dancing Raisins scenario was randomly selected from within 10 freshman students' maps, one from within 10 teachers, and one from within 10 scientists. Similar random selection was made for the remaining three scenarios from the three participating groups. This selection ensured that the two researchers together analyzed explanations maps of all four scenarios from all three groups. Figure 3.3 presents an overview of the study's participant groups, time-line, procedures, instruments, and data sources.

First, the researcher introduced the scholar to the NOSE framework, the procedures involved in building explanation maps, and the process by which these maps were analyzed using the NOSE framework. Next, the scholar received a randomly chosen segment of an interview transcript of one of the four scenarios in interview I, which included one final explanation of this scenario generated by a participant along with its corresponding un-coded explanation map (i.e., all statements were in concept balloons). The identity and group membership (i.e., student, teacher, or scientist) of the participant transcript used was not shared with the scholar. The two researchers (the primary researcher and the scholar) read the corresponding transcribed segment, had the corresponding explanation map and then analyzed the final explanation of the scenario. The two researchers then met to discuss their analysis. Discussions focused on the coding of structural elements of the statements and how the map captured the participant's explanation through the NOSE framework. Furthermore, discussions emphasized the extent to which NOSE framework accurately depicted the participant's explanation. The analysis was used to identify the nature (structural elements) and quality (completeness and adequacy) of participants' final explanations in addition to identifying the type(s) of explanations present. Discussion went on until the two researchers reached a consensus analysis of the explanation map. Throughout this process, the researcher acted as a

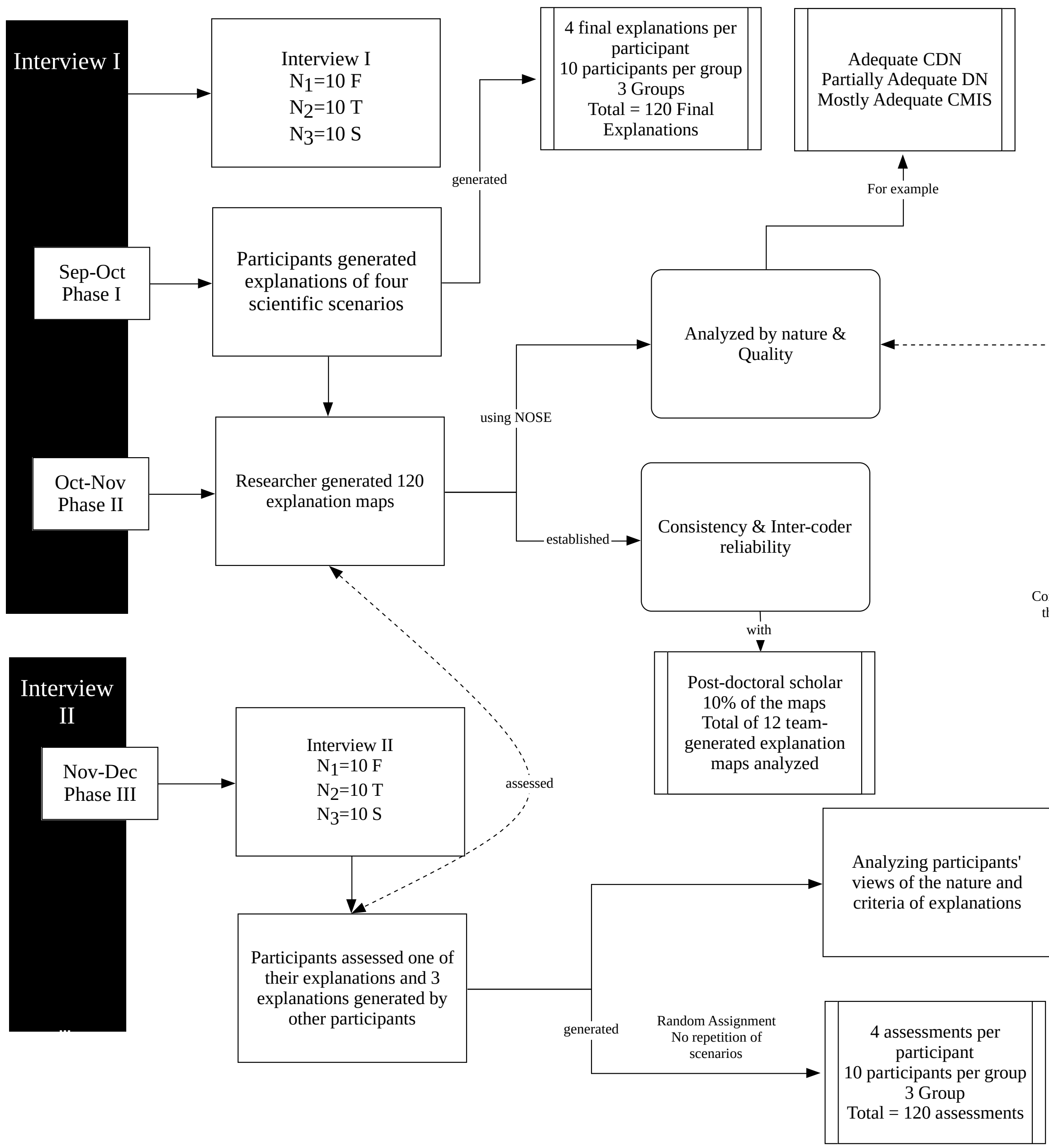


Figure 3.3 An overview of the study' participant groups, timeline, instruments, and data sources.

facilitator and avoided influencing the explanation analysis. The researcher only clarified statement(s) in the transcripts that were not clear and provided relevant contextual information.

The researcher then provided the scholar with 10% of explanations maps of the final explanations from each group of participants. The identities and group membership (i.e., student, teacher, or scientist) of participant transcripts was not shared with the scholar. The researchers independently analyzed corresponding explanation maps of final explanations according to the NOSE framework and following the consensus reached regarding the criteria to use when analyzing explanation maps. They then met on weekly basis to exchange and discuss individual map analysis that they conducted for each explanation map. Discussions mostly focused on comparing the analysis of the explanation maps and analyzing the general direction of the proposed explanation. Disagreements were resolved through further discussion until consensus was achieved on a final explanation map analysis of each participant's transcript of the corresponding final explanation.

Next, the researcher constructed explanation maps of the remaining 96 explanation maps of the final explanations generated by all three participating groups. The analysis followed NOSE framework and the agreed upon criteria reached between the two researchers. The analysis was used to identify the nature (structural elements) and quality (completeness and adequacy) of participants' final explanations in addition to identifying the type(s) of explanations present. For a more detailed discussion of the NOSE framework, see Part 3 of Chapter 2.

Finally, the researcher examined the characterizations of the explanation maps for each group of participants to generate a full descriptive profile of these maps. Each profile detailed the characteristics of a participant group's maps. The profiles were compared and contrasted both within and across participant groups and assertions regarding ways in which students, teachers, and scientists' explanations were similar or different in accordance with the NOSE framework were made.

Part II. Analyzing participants' views of the nature and criteria of explanations. A major purpose of this study was to examine the extent to which the criteria for assessing scientific explanations generated by participants aligned with those in the NOSE framework. Thus, albeit the analysis included all three groups of participants, of particular interest were the criteria used by the practicing scientists given the reforms' emphasis on the need for attaining instructional outcomes for science students within authentic scientific practice (e.g., NRC, 2000). Another purpose of the study was to assess whether or not NOSE, a formal analytical framework guided by philosophical models of explanations, placed realistic expectations on students' construction and assessment of scientific explanations. This part of data analysis focused on comparing and contrasting the criteria derived from the first part of analysis with those in the NOSE framework.

Hence, after conducting the first round of interviews (Interview I) and after generating explanation maps of participants' final explanations, corresponding transcripts in addition to the explanation maps were used to conduct Interview II. Transcripts generated during the second interview were used to characterize participants' views of the nature of scientific explanation, and derive the criteria used by members of the three groups to judge the nature and quality of explanations. This latter analysis was followed by comparing and contrasting the analyses within and across the three groups. Furthermore, since the generated explanation maps were analyzed according to the NOSE framework, the analysis focused on comparing the criteria derived from the participants with aspects of explanation emphasized in the NOSE framework.

Thus, this part of data analysis focused on analyzing transcripts produced from Interview II where participants assessed and provided feedback on the final explanations generated during Interview I by (1) themselves, (2) participants in their own group, and (3) participants from the other two groups. Participants' views of the nature and criteria used to judge the completeness and adequacy of explanations were analyzed and individual profiles of the views and nature of explanation and the criteria used to judge explanations were generated. Profiles within each group of participants were examined for general patterns in order to produce a set of criteria that each group used in their assessment. The generated sets were compared and contrasted across the groups in an aim to answer the fourth research question.

Limitations of the Study

First, because of the self-selected nature of participants, this study does not claim to derive generalizable results. The participants were not necessarily representative of a larger group of freshmen college students, science teacher, and practicing scientists. Nonetheless, the results obtained were valuable in shedding light on the appropriateness of the expectations for using NOSE framework to examine scientific explanations. Second, participants' content knowledge of the scientific concepts addressed in this study constituted a confounding factor that could not be controlled for. Participants' prior knowledge related to these concepts affected both their explanations and the criteria they generated to judge the goodness of the explanations of others. During the first round of interviews, participants were asked whether or not they were familiar with the phenomena at hand. While it did not constitute a full-blown assessment of their prior knowledge, participants' self-reported datum was factored into the analysis. Third, there exists some circularity in the design of this study: because philosophical models of explanation constitute robust support in the construction of a framework unique to explanation, the NOSE framework was used to analyze participants' explanations. At the same time, participants' views of explanation and the criteria they provided regarding the completeness and quality of explanations played a vital role in assessing the practical validity of the NOSE framework itself.

CHAPTER 4: RESULTS

The first section of this chapter presents an analysis of participants' scientific explanations using the NOSE framework. The second section explicates participants' perceptions of explanation, and the third, the criteria deployed by participants to assess the "goodness" of explanations. The fourth and final section discusses how the criteria used by participant groups compare to those underlying the NOSE framework. In the following sections, pseudonyms are used to refer to participants. Freshman students pseudonyms will begin with the letter F, teachers with T, and scientists with S.

During the first interview, all participants generated an explanation in relation to four scientific phenomena. Explanation maps were then developed to visualize participants' explanations of the "why question" posed by the researcher. The following sections report both major (50% and more occurrence within each group of participants) and minor features (20%-40% occurrence within each group) that were evident in participants' explanation maps.

Analysis of Participants' Explanations

For each scenario, explanation maps were constructed using participants' verbatim transcripts of their explanations generated during interview I. The present study analyzes in-depth participants' explanations from the first two scenarios. Both CIJ and DR scenarios involved simple materials but included complex scientific understanding. The two scenarios elicited a large number of scientific explanations that provided a rich context for analysis. For the first scenario, the Dancing Raisins (denoted by DR in this chapter), each participant's explanation was an answer to "Why did the raisins first sink to the bottom, and why did they then float up to the top, then sink again?" In the second scenario, the Candle in the Jar (denoted by CIJ in this chapter), each participant's explanation was an answer to "Why did the water rise when the lit candle was covered with an inverted glass jar?"

Each explanation map was constructed from a participant's coded transcript using the NOSE framework procedures. A detailed description of the construction of explanation maps is presented in Chapter 3. The structural elements identified in the maps aimed to represent participants' type, as well as nature and quality, of their generated explanations of each phenomena (see Table 3.1). In both scenarios, common patterns were observed across all participants' maps in relation to various structural elements. The following sections present results from participants' explanations of the first two scenarios (CIJ and DR scenarios).

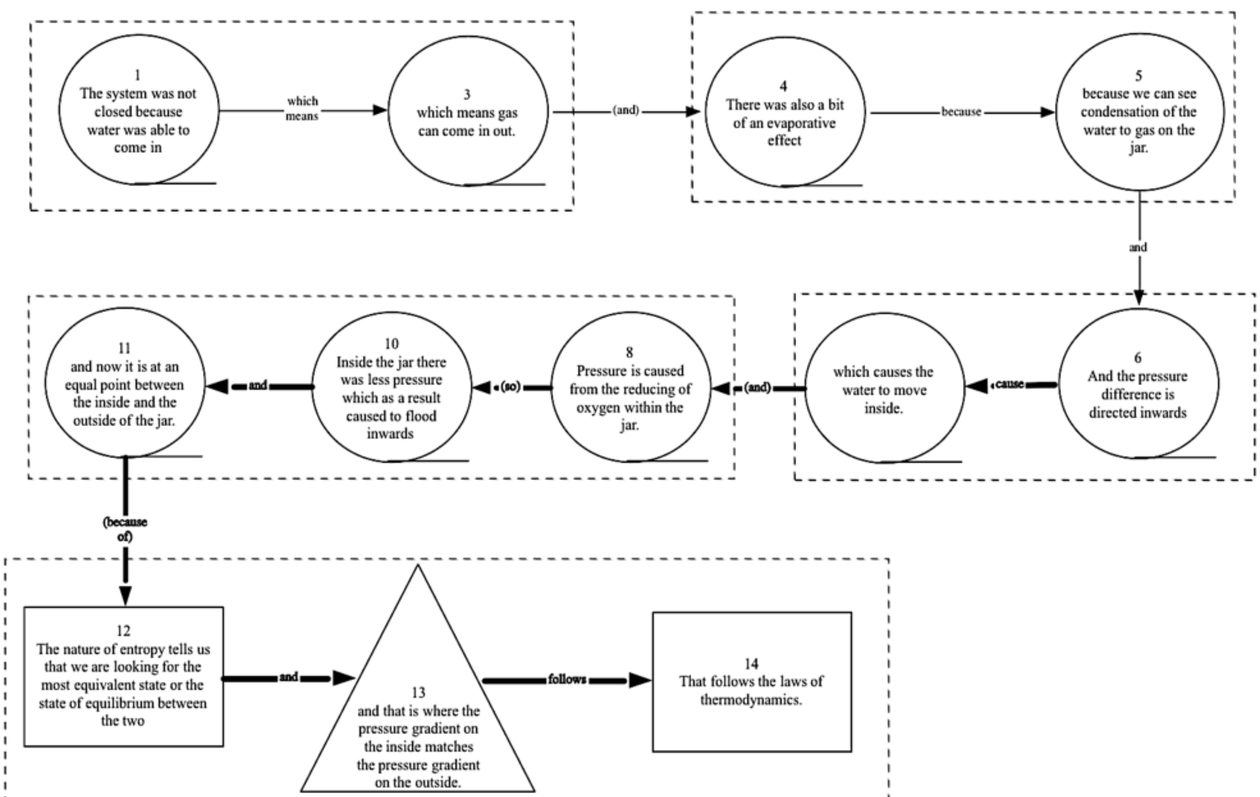


Figure 4.1 Sample explanation map of the CIJ scenario constructed by Tucker, a participant teacher. All shapes contain statements directly excerpted from participant transcripts. Numbers are used to identify the sequence of statements. Different shapes indicate different structural elements as presented in Table 3.1. Arrows show the directional flow of statements for a participant. Bolded arrows indicate a causal link. Dotted rectangles containing several structural elements indicate one explanatory idea.

The Candle in the Jar (CIJ) Scenario

Basic features of explanation maps across groups. In the CIJ scenario, results from participants' explanations of the water rising phenomenon revealed common patterns across all participants in relation to the use of observations and inferences versus pieces of knowledge, laws and lawlike statements, the use of teleological and anthropomorphic statements, and the nature of explanatory connections or big ideas made. Each of these aspects is discussed below.

Observations and inferences. According to the NOSE framework, scientific observations are descriptions of phenomena through the senses or extensions of the senses; whereas scientific inferences are interpretations based on these observations, which are not directly available to the senses (Lederman, 2007). In explanation maps, statements contained in concept balloons are observation statements related to the phenomenon-to-be-explained; while statements contained in concept balloons with a horizontal line at the bottom are inferences (see Table 3.1).

During the first round of interviews, participants were asked to describe what happened when the lit candle was covered with an inverted glass jar and to explain why the water rose in the jar. Table 4.1 shows that for this phenomenon, students and teachers demonstrated the highest use of observations and inferences compared to scientists: out of a total of 62 observations generated by all participants, 45% (28 observations) were produced by students and 35.5% (22 observations) by teachers as compared to 19.3% (12 observations) produced by scientists. Similarly, out of 140 total inferences generated by all participants, 41% (57 inferences) were produced by students and 39% (54 inferences) by teachers as compared to 21% (29 inferences) produced by scientists.

Aligned with the NOSE framework, observations and inferences were assessed based on their relevance and accuracy. Two types of observation statements in relation to CIJ were observed: accurate relevant observations and inaccurate relevant observations. On the other hand, three types of inference statements were observed among, at least, two of the three groups: accurate relevant inferences, inaccurate relevant inferences, and accurate irrelevant inferences.

Table 4.1

Frequency of Occurrence of Major and Minor Structural Elements Used by Participants in Explanation Maps of the CIJ Scenario Across the Three Groups

	Scientists	**Teachers**	**Freshman**	**Total**
Structural Elements				
Observation	12	22	28	**62**
Inference	29	54	57	**140**
Piece of Knowledge	38	25	11	**74**
Necessary Condition	14	9	9	**32**
Laws & Lawlike Statement	24	13	9	**46**
Teleological/Anthropomorphic statement	4	9	14	**27**

As evident in Figure 4.2, all participants' observation statements were relevant to the CIJ phenomenon; however, some of observations were accurate, while others were inaccurate. Recall that scientists produced the least number of total observations among the three participating groups. Additionally, they were the only group that did not produce any inaccurate observations (i.e., all scientist-produced observation statements of the CIJ scenario were accurate and relevant). On the other hand, 8 of the 28 observation statements produced by students were inaccurate as compared to two inaccurate observations produced by teachers.

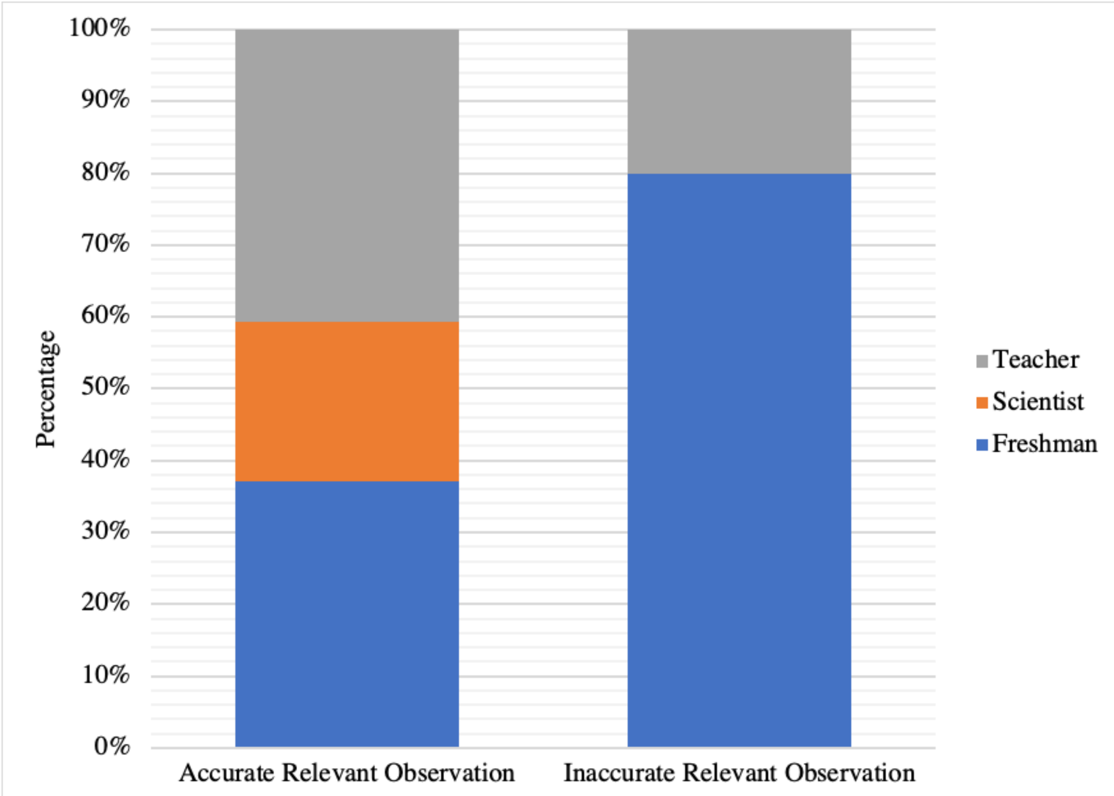


Figure 4.2 Participants' types of observations in the CIJ scenario. The types are based on percentage frequencies. Note that Inaccurate Relevant Observations on the X-axis is a type of observation statements that was observed in only two participant groups.

An interesting finding revealed that inaccurate observations were majorly related to the bubbles observed at the end of the CIJ demonstration. While only five students were able to observe the bubbles during the CIJ demonstration, all of them made inaccurate observations related to the bubbles. In particular, 4 of the 5 students said that they observed air getting *into* the jar instead of escaping it. For example, Finn stated that "air started to go in through the opening," and Fredrick said: "and then as soon as liquid got in, some air started getting in." Another inaccurate observation that was not related to the bubbles was stated by Florence: "The water rises up until none of the water is outside the jar"; even though this was not the case when she observed the demonstration with different jar and candle sizes. Other inaccurate observations were related to the sequence of events of CIJ. In particular, Finn claimed to observe the flame going out when the water finished rising, while Felicia claimed observing that "once the bubbles started, the flame was put out." Finally, two teachers provided inaccurate observations also related to the sequence of events in the CIJ scenario: both Tammy and Tarra stated that they observed the water stopped rising as soon as the flame went out.

As mentioned earlier, three types of inference statements related to the CIJ demonstration were observed: accurate relevant, accurate irrelevant, and inaccurate relevant inferences. Similar to findings from participants' observations, scientists produced significantly less inference statements than teachers and students. As evident in Figure 4.3, scientists were the only participating group that did not produce any irrelevant inferences. While the majority of inferences produced by all participants were accurate and relevant, freshman students produced significantly more irrelevant and inaccurate inferences than teachers and scientists.

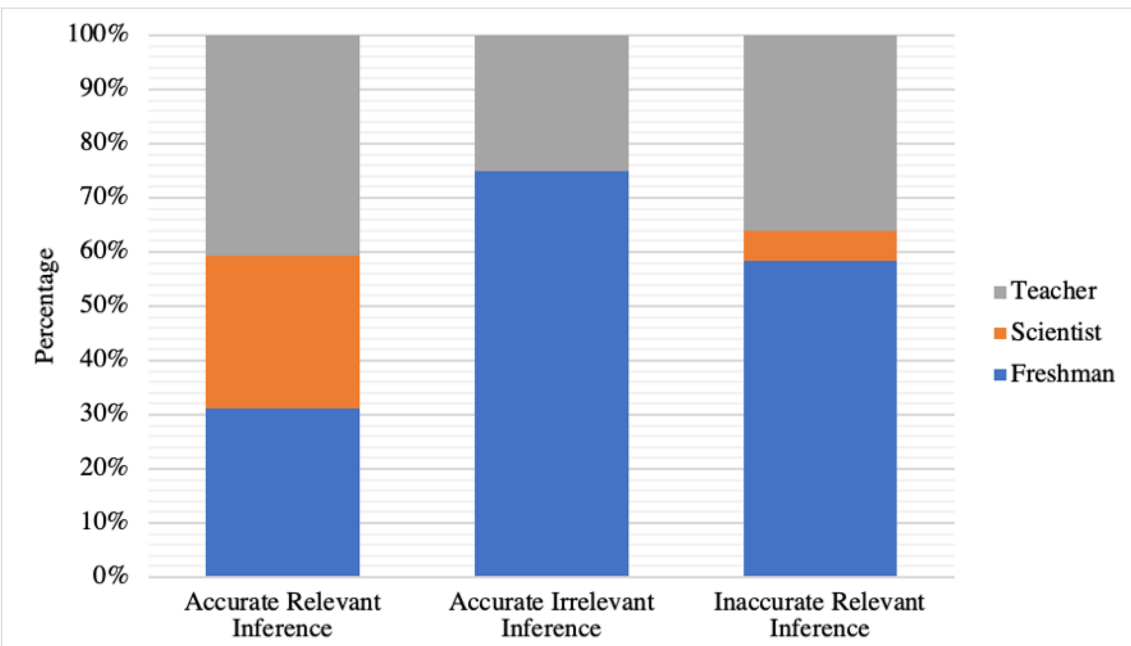


Figure 4.3 Participants' types of inferences produced in the CIJ scenario. The types are based on percentage frequencies. Note that Accurate Irrelevant Inference on the X-axis is a type of inference statements that was observed in only two participant groups.

In particular, 6 irrelevant inferences (75%) of all accurate irrelevant inferences were produced by students as compared to 2 irrelevant inferences (25%) by teachers. Additionally, 21 inaccurate inferences (58.3%) of all relevant inaccurate inferences were produced by students as compared to 13 inaccurate inferences (36.2%) produced by teachers, and only 2 inaccurate inferences (5.5%) by scientists.

Another finding reveals that, in general, inferences included common ideas related to the CIJ demonstration across the three groups. In particular, all participants stated that the flame went out because there was a loss of oxygen inside the jar. Some participants further added a necessary condition that this happened after the candle was covered by the jar. For example, Fidel stated: "As you are covering it [the candle], you are restricting the air around the candle to just what is inside the jar." A more important focus of this study is related to the water rising phenomenon in the CIJ demonstration (i.e., an explanation of why the water rose). The majority of participants' inferences in this regard included inferences about *pulling*, *pushing*, or even *sucking* of water, inferences related to consumption and burning of oxygen, and lastly inferences related to decreasing pressure and pressure equilibrium.

The water was "pulled," "pushed," or "sucked" inferences. While describing the water behavior when the candle was covered, participants from the three groups tended to use inferential terms such as "pull," "push," and "suck" to describe the water rising. For example, Stanley stated that "eventually the water was *pulled* [emphasis added] up through the jar"; Fredrick said that "the water started to *push* [emphasis added] up"; and Todd stated that "the water is being *sucked* [emphasis added] into the canister." It is worth nothing that inferences describing the water rising as being sucked or the candle pulling the water in were considered inaccurate inferences. What is more, significantly more students and teachers used these terms than scientists; scientists tended to produce more observation statements and used terms such as "the water level rises more" (Sam) or "the water goes up" (Samantha). Other scientists were even more specific in describing the behavior of the water after it was covered by the jar. For example, Stefan stated that "the water level does not rise instantly but continues to rise after the flame goes out."

The consumption and burning of oxygen inferences. When explaining the water rising phenomenon in the CIJ demonstration, another common inference produced by participants from the three groups was related to the oxygen inside the jar. Four students stated that when the candle was covered, oxygen was used up or became limited. For example, Filip stated that "when it [the water] went inside, the oxygen that was inside the water was used up"; while Flynn was more specific in his inference as he explained: "When you put a jar on top of a candle that is burning, now there is a limited amount of air that it can burn through, specifically oxygen." Oxygen-related inferences were more common among scientists and teachers than students. In particular, six scientists and five teachers produced inferences related to the burning and consumption of oxygen inside the jar. For example, Sara said that "when you burn a candle in an enclosed space it is going to burn up all of the oxygen within the system"; while Tammy stated that "the water goes up because the oxygen was depleted."

Pressure-related inferences. The final major common inference that was observed among the three participating groups was one related to pressure. After observing the water level rising inside the jar, 4 of 10 students, 7 of 10 teachers, and 9 of 10 scientists produced inferences related to decreased pressure. Students and teachers tended to support their inferences with their observations or with nothing at all; whereas scientists supported their inferences with pieces of knowledge and lawlike statements. For instance, Finn stated that "the water was pulled in due to decrease in air pressure inside the jar," and Fredrick elaborated: "Inside there is some kind of vacuum or something that is leading to very low pressure, which is why the water came up, because there is low pressure inside." Similarly, after stating that the water was pulled into the jar, Tucker further said that "there is a pressure difference within the jar and outside of the jar."

While scientists produced similar inferences, they added some pieces of their prior knowledge (i.e., pieces of knowledge) and lawlike statements to support the inferences they made. For example, consider Sara's CIJ explanation in the following excerpt and its corresponding explanation map (Figure 4.4):

> When you burn a candle in an enclosed space it is going to burn up all of the oxygen within the system and then you are going to have a bunch of free space that previously had something in it. And that creates what's called a vacuum, which causes a difference in pressure between the outside and the inside. That means that outside the jar we are sitting under atmospheric pressure with all the air pushing down on us but inside the jar by burning the candle you have essentially removed a portion of that. So, there is not as much pressure pushing down on the system inside the jar as there is on us or the water outside the jar. So, when the candle burns it up, the water goes in to fill up the void space because the pressure pushing down on the outside is greater than the pressure inside, so it is going away from where the high pressure is into where the low pressure. This is Le Chatelier Principle.

Sara's explanation of why the water rose is not fully adequate as she did not take into account the production of carbon dioxide and water vapor. However, in her explanation, Sara tended to support her inferences with pieces of knowledge pertaining to atmospheric pressure, vacuum, and pressure differential, and subsumed her explanation of the phenomenon under Le

Chatelier principle. Another scientist example is that of Samantha who even though acknowledged that when oxygen was consumed another gas was produced, they were not the same amount. Samantha, however, did not provide any further clarification. Part of Samantha's CIJ explanation is presented in the following excerpt with its corresponding explanation map (Figure 4.5):

> When we cover the candle, we are diminishing the amount of oxygen in the jar. Upon burning the oxygen even if there is some gas that is created it is not the same amount as there was before as was with oxygen. There is some loss of gas, so there is less pressure inside, so there is an imbalance between the pressure inside and the pressure outside so the water goes up to equalize the pressure.

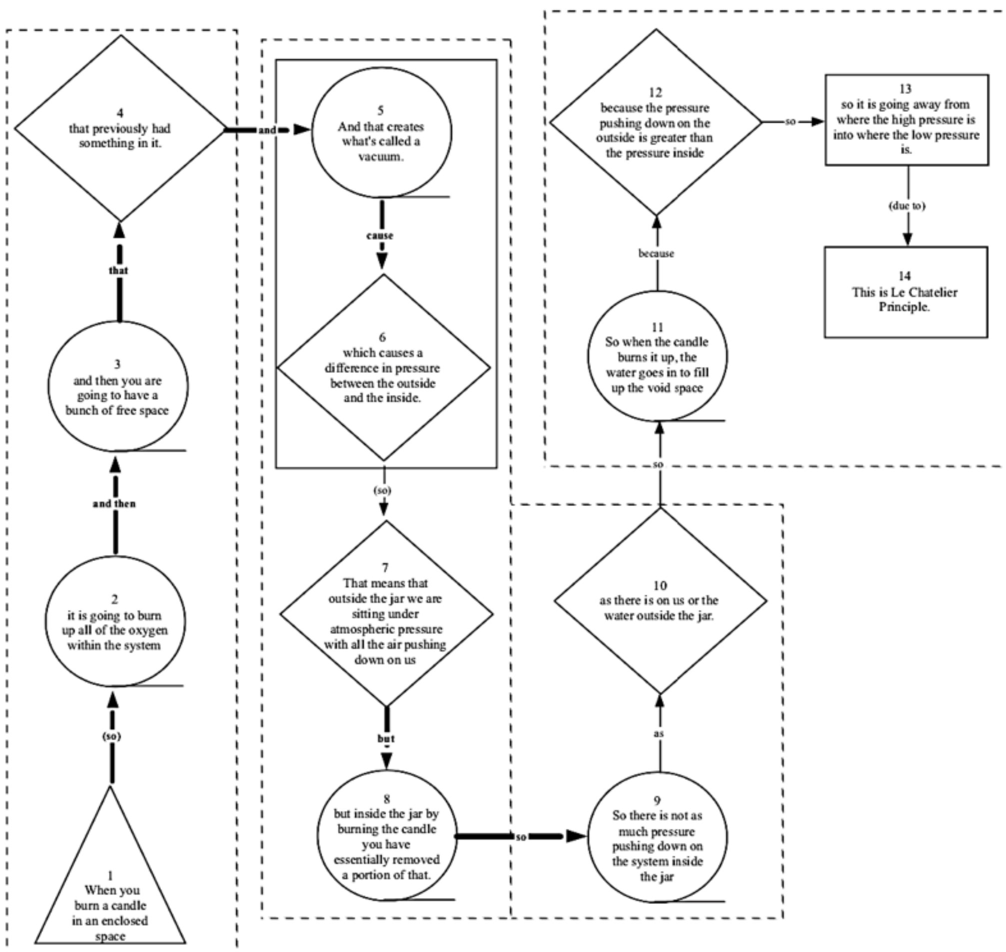


Figure 4.4 Sara's explanation map. An example of a scientist's map of CIJ explanation.

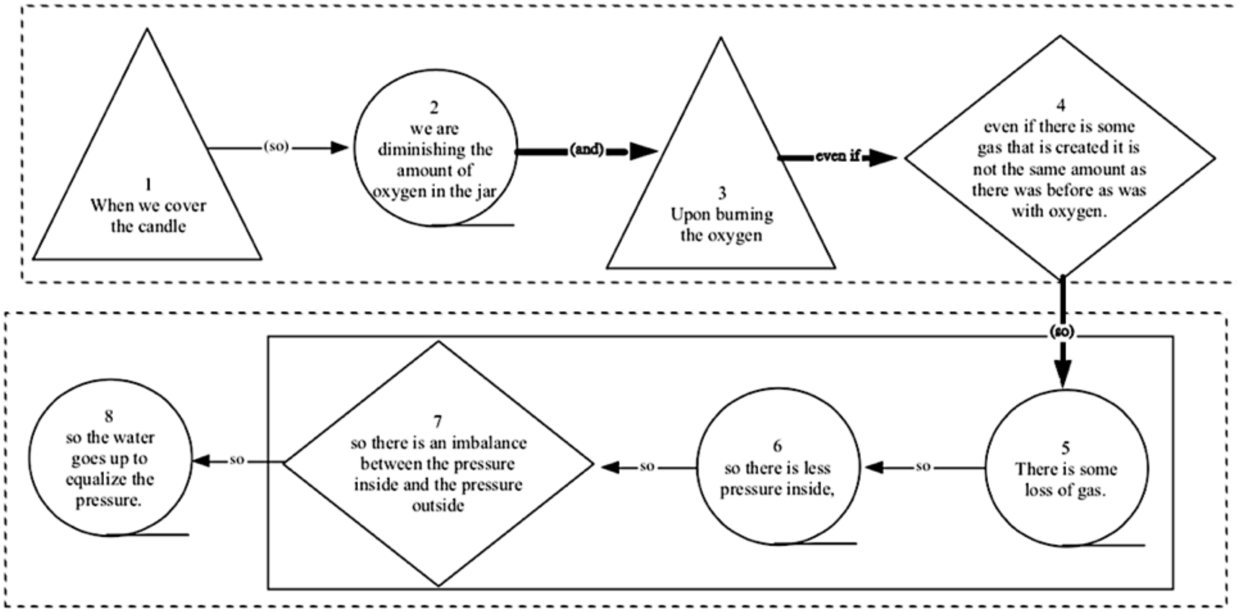


Figure 4.5 A part of Samantha's explanation map. An example of a scientist's map of the CIJ explanation.

Pieces of knowledge (PK). According to the NOSE framework, pieces of knowledge refer to previously learned information, or prior bits of knowledge. Diamond shapes in explanation maps contain pieces of knowledge (see Table 3.1 for more details). As with observations and inferences, pieces of knowledge are assessed based on their relevance and accuracy. As evident in Table 4.1, scientists demonstrated the highest use of pieces of knowledge: out of a total of 74 pieces of knowledge generated by all participants, 51.3% (38 PK) were produced by scientists, 34% (25 PK) by teachers, and 15% (11 PK) by students.

As mentioned earlier, while students and teachers supported their inferences with their observations of the phenomenon, scientists tended to use significantly more PK to support their inferences and observations. Hence, it is meaningful to examine the percentage frequencies of observations, inferences and pieces of knowledge across the three groups. As evident in Figure 4.6, students and teachers produced significantly more inferences and observations than scientists: 45% of all observations were produced by students, 35% by teachers and 19% by scientists. Additionally, 41% of all inferences were produced by students, 38.5% by teachers, and 27% by scientists. However, scientists produced significantly more pieces of knowledge than teachers and students: 51.3% of all PK were produced by scientists, 39% by teachers, and only 16% by students.

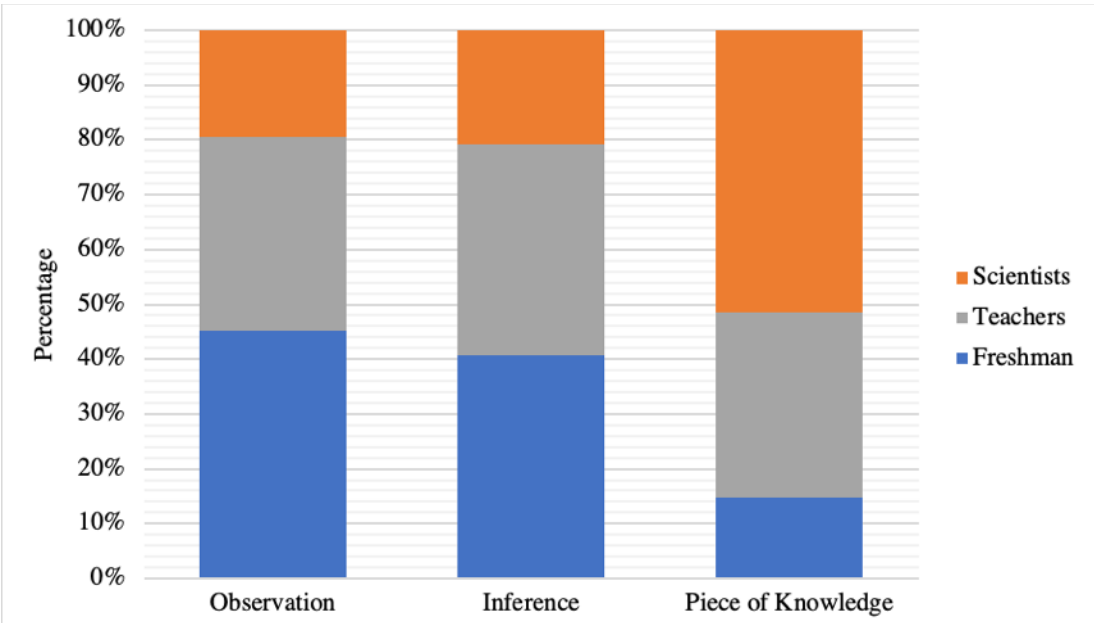


Figure 4.6 Participants' percentage frequencies of observations, inferences and pieces of knowledge in the CIJ demonstration across the three groups.

Additionally**,** three types of PK were observed: accurate relevant PK, accurate irrelevant PK, and inaccurate relevant PK. Of the 38 PK produced by scientists, 8% were irrelevant accurate, and only 3% were inaccurate relevant (the remaining 89% were accurate relevant PK). However, 20% of the 25 PK produced by teachers were inaccurate relevant PK and the remaining were accurate relevant. Finally, while students produced only 11 PK's, 54.5% of them were accurate relevant, 36.3% of them were irrelevant accurate, and the remaining one PK was inaccurate relevant PK.

Further examination of Figure 4.7 reveals that not only did scientists produce the highest number of pieces of knowledge in their CIJ explanations, they also produced the highest percentage of accurate PK as compared to students and teachers. In particular, 34 accurate relevant PK (57%) of all accurate relevant PK were produced by scientists, 20 (33%) by teachers, and only 6 (10%) were produced by students.

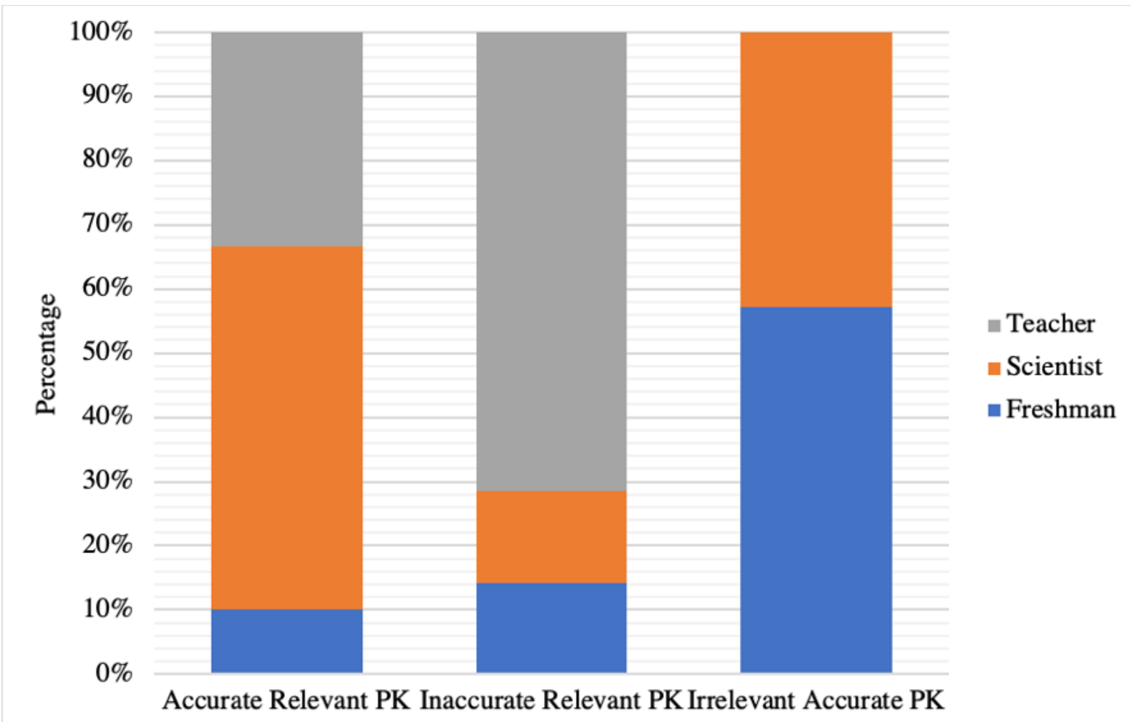


Figure 4.7 Participants' types of pieces of knowledge (PK) in the CIJ demonstration. The types are based on percentage frequencies. Note that Irrelevant Accurate PK on the X-axis is a type of inference statements that was observed in only two participant groups.

Laws and lawlike statements. According to the NOSE framework, there are two main types of laws: general laws and lawlike statements and probabilistic-statistical laws and lawlike statements. Solid rectangular shapes in explanation maps contain lawlike statements of general laws and natural regularities; whereas dashed rectangular shapes contain statements of laws of probabilistic and statistical nature (see Table 3.1). These lawlike statements are not necessarily accurate, logical, or canonical. Results show that the vast majority of laws and lawlike statements used in CIJ explanations produced by all participants were general laws and natural regularities: 43 of 46 lawlike statements were of deterministic nature, while only three were of a probabilistic-statistical nature produced by one participant scientist only. In particular, Sam explaining the pressure difference in the CIJ demonstration stated (see Figure 4.8):

> This is a probability related thing. This does not have to happen since the molecules are bouncing around randomly. You could have a situation where all the molecules on the inside just randomly bounce downwards and push all the water out, but that is extremely unlikely.

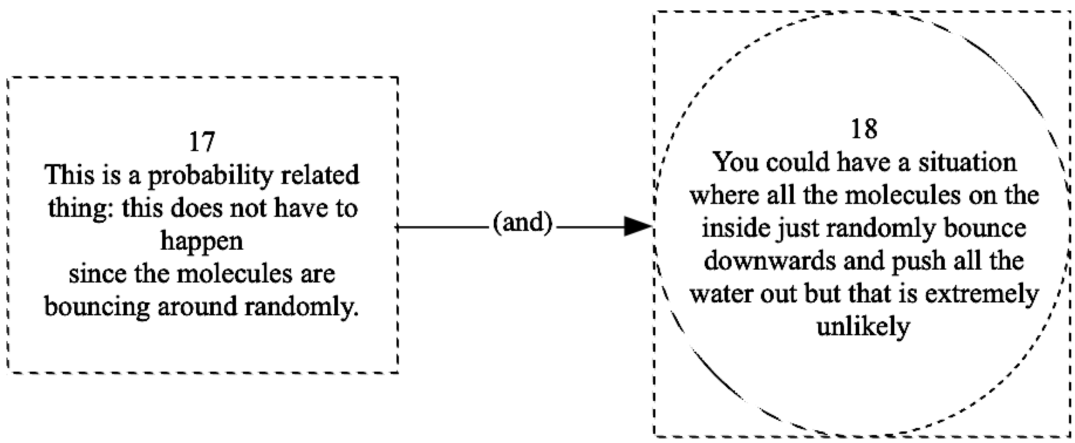


Figure 4.8 A part of Sam's explanation map. An example of a scientist's map showing a probabilistic lawlike statement and a probabilistic prediction.

It is worth noting that even though Sam used probabilistic/statistical-type statements, they are not necessarily accurate. Sam stated that it was extremely unlikely for the molecules inside the jar to push the water out, but he did not further explain why it was unlikely.

On the other hand, Tucker referred to the laws of thermodynamics while providing his CIJ explanation (see Figure 4.1, structural element numbers 12,13, 14):

The nature of entropy tells us that we are looking for the most equivalent state or the state of equilibrium between the two and that is where the pressure gradient on the inside matches the pressure gradient on the outside. That follows the laws of thermodynamics.

As would be expected, scientists demonstrated the highest use of lawlike statements compared to teachers and students: 24 of 46 lawlike statements were produced by scientists as compared to 13 produced by teachers and 9 by students. Further examination of the results reveal that only half of participating students produced lawlike statements as compared to 9 of 10 scientists and 8 of 10 teachers. Additionally, out of the five students who used general laws and lawlike statements, three explained the water rising phenomenon as a consequence of pressure only (pressure difference, pressure drop, or pressure equilibrium), while the remaining two students explained the phenomenon as a consequence of some irrelevant lawlike statements. In particular, Franco explained the water rising due to water potential; while Fidel explained the phenomenon as due density of gases. On the other hand, the vast majority of participating scientists and teachers who used laws and lawlike statements, referred to more than one law or lawlike statement to explain the water rising phenomenon.

Table 4.2 shows the different laws and lawlike statements that participants used to explain the water rising phenomenon. Unsurprisingly, two common laws and lawlike statements used in participants' explanations of the water rising phenomenon were mainly related to consumption of oxygen and/or pressure—which is essentially not adequate. Three freshman students explained the water rising phenomenon as a logical consequence of pressure difference without including any other pieces of knowledge or necessary conditions, and without further explaining the pressure difference. For instance, below is a part of Filip's CIJ explanation (also see Figure 4.9):

The pressure inside decreased. The water was under atmospheric pressure, but as soon as the candle started burning, it decreased the pressure in some way. In nature there is tendency for things to go from high pressure to low pressure.

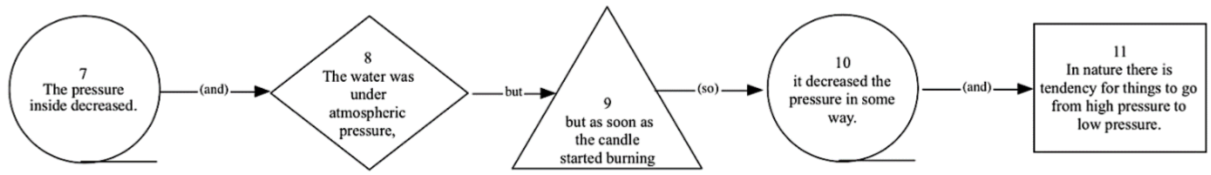


Figure 4.9 A part of Filip's explanation map. An example of a student's map showing the explanation of CIJ scenario as a consequence of pressure.

On the other hand, the majority of participant teachers explained the CIJ phenomenon as a consequence of the pressure difference, which was in turn due to other lawlike statements and related pieces of knowledge and necessary conditions. Similarly, no participant scientist explained the water rising phenomenon due to pressure only; instead, when they explained it as a logical consequence of pressure (pressure drop or pressure difference), they further explained the pressure decrease (or pressure difference) using other lawlike statements, pieces of knowledge and necessary conditions. However, there were still differences in the use of these laws and lawlike statements between teachers and scientists: teachers who explained the water rising phenomenon due to pressure decrease in the jar further explained the pressure decrease as a result of oxygen consumption. On the other hand, most scientists who explained the water rising phenomenon due to pressure decreasing in the jar, further explained that even though oxygen was consumed inside the jar, other gases were produced, and therefore added other pieces of knowledge to explain how even with the production of other gases, the consumption of oxygen still lead to pressure decrease. It is worth noting that they still did not adequately explain the CIJ phenomenon (an example of an adequate explanation of the CIJ phenomenon is presented in Figure 4.12). Recall in Figure 4.5 Samantha explained that "even if there is some gas that is created it is not the same amount", and Selena explained that "the partial pressure of oxygen is more than that of carbon dioxide".

Table 4.2

The Use of Laws and Lawlike Statements in Explaining the Water Rising Phenomenon Among All Participants (All names are pseudonyms)

	Pressure	Oxygen Consumption	Temperature Change	Expansion/ Contraction	Ideal Gas Law	Other
Finn F1	✓	--	--	--	--	--
Felicia F2	--	--	--	--	--	--
Faith F3	--	--	--	--	--	--
Florence F4	--	--	--	--	--	--
Farrah F5	--	--	--	--	--	--
Flynn F6	--	--	--	--	--	--
Franco F7	--	--	--	--	--	Water Potential
Filip F8	✓	--	--	--	--	--
Fidel F9	--	--	--	--	--	Density
Fredrick F10	✓	--	--	--	--	--
ThomasT1	✓	✓	✓	✓	✓	Heat
Tucker T2	✓	✓	--	--	--	Entropy, thermodynamics
Trevor T3	✓	--	✓	--	--	Gas condensation
Tina T4	✓	✓	--	--	--	Fluids behavior
Tammy T5	✓	✓	--	--	--	--
Tanya T6	--	--	--	--	--	--
Tarra T7	--	--	--	--	--	--
Tod T8	✓	✓	✓	✓	✓	--
Tracy T9	✓	--	✓	--	--	Heat
Tyson T10	--	--	--	--	--	Heat
Sara S1	✓	✓	--	--	--	Le Chatelier
Sophia S2	✓	✓	--	--	--	--
Sam S3	✓	--	✓	--	✓	--
SamanthaS4	✓	✓	--	--	--	--
Selena S5	✓	--	✓	--	✓	--
Stanley S6	✓	✓	--	--	--	Speed of Gases
Stefan S7	✓	✓	✓	✓	✓	--
Stella S8	--	✓	--	✓	--	Osmosis
Sylvia S9	--	--	--	--	--	--
Saul S10	✓	✓	✓	✓	✓	--

As evident in Table 4.2, five teachers and six scientists included laws and lawlike statements pertaining to pressure and oxygen consumption to explain the water rising phenomenon (which, it should be noted is not a scientifically adequate explanation). However, 2 of the 5 teachers and 4 of the 6 scientists used these two lawlike statements in addition to other laws, lawlike statements and pieces of knowledge. For instance, Thomas explained the water rising phenomenon due to pressure difference and equilibrium, oxygen consumption, temperature change, expansion and contraction of the gas, and ideal gas law. On the other hand, Stanley explained the water rising phenomenon due to pressure, oxygen consumption and the speed of gases inside the jar – where the latter was an irrelevant piece of knowledge. Finally, temperature change, expansion and contraction, and Ideal Gas Law were three laws that were used by four teachers and five scientists. In particular, two of these four teachers explained the water rising phenomenon according to pressure and temperature change, while the other two included lawlike statements related to the expansion and contraction of gases associated with the temperature and pressure changes.

The use of teleological and anthropomorphic statements. According to the NOSE framework, teleological statements refer to something happening as a function of its end, purpose, or goal. Additionally, anthropomorphic statements include statements that ascribe human feelings and behaviors to elements of the phenomenon-to-be-explained. Recall that these statements are included in hexagon shapes in explanation maps (see Table 3.1 for more details). As evident in Table 4.1, participants from the three groups tended to use, though with varying degrees, these statement types while generating their CIJ explanations. In particular, freshman students demonstrated the highest use of these statements: out of a total of 27 teleological/anthropomorphic statements, 14 (52%) were produced by 3 students, 9 (33.3%) were produced 3 teachers, and 4 (15%) teleological/anthropomorphic statements were produced by only one scientist. For example, while explaining the CIJ phenomenon, Felicia, Florence, and Farrah heavily used anthropomorphic and teleological statements, such as “the water was trying to rise up to get to the flame” (Felicia), “the oxygen in the water will want to feed the flame. And that’s why the water keeps coming up into the jar. And so the flame is attracted to oxygen, and wants to be fed by oxygen” (Florence), and “the heat from the candle tried to escape the jar. We also heard the grr sound; it was the water trying to rush in the jar” (Farrah). It is worth noting that—and this is elaborated later in this chapter–these three students did not produce scientific explanations of the CIJ phenomenon according to the NOSE framework. Additionally, Sylvia, the one participating scientist, used similar anthropomorphic and teleological statements while constructing also a non-scientific explanation according to the NOSE framework. The following is an excerpt of Sylvia’s explanation of the CIJ phenomenon followed by its corresponding explanation map (Figure 4.10):

> As the flame needed more oxygen, it started bringing the water into the jar until all the oxygen was used up. As the candle uses up the oxygen it keeps pulling to get all the oxygen it can, and then when it goes out, there is still that force remaining to pull all the oxygen from the water in.

An examination of the three teachers’ explanations who used anthropomorphic and teleological statements reveals different results. In accordance with the NOSE framework, two (Tracy and Tyson) out of these three teachers still produced scientific explanations (though only partially adequate) that included anthropomorphic and teleological statements, while the remaining teacher produced a non-scientific explanation. It was evident that in Tracy’s and Tyson’s explanations – the two teachers who despite using anthropomorphic statements produced scientific explanations – these statements were tightly linked to pieces of knowledge and laws of which the CIJ phenomenon was a consequence.

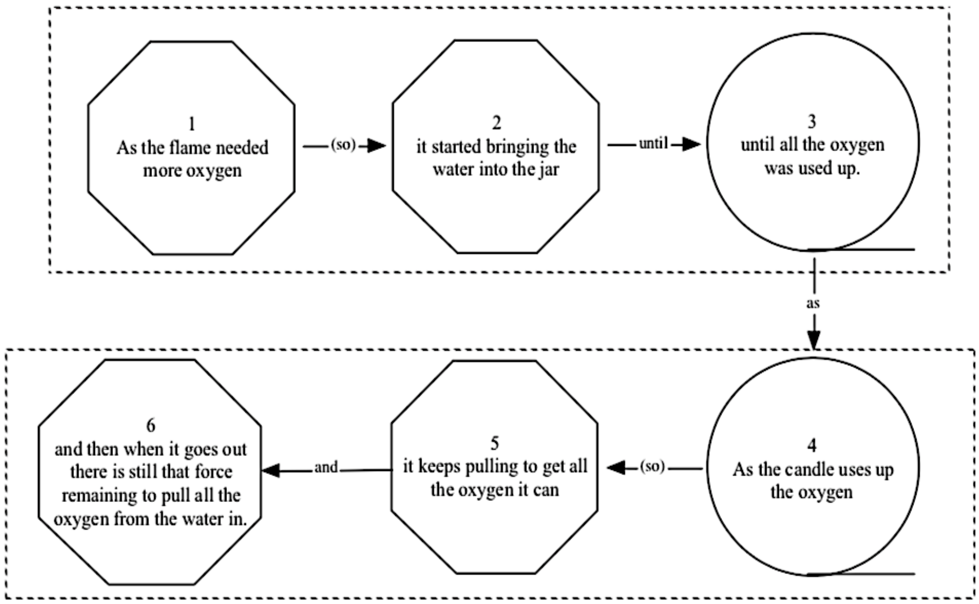


Figure 4.10 Sylvia's explanation map. An example of a scientist's map showing teleological and anthropomorphic statements in her non-scientific explanation.

For example, consider Tracy's explanation and its corresponding map (Figure 4.11):

> The water was pulled into the system when the air was trapped inside. Not only are you changing the temperature inside, but you are also changing the pressure inside. The water is pulled in to equalize the pressure because heat is related to pressure, and the gases that are in there are at a different pressure now. *The system wants to be in equilibrium – most things throughout our world want to be in a state of equilibrium* [emphasis added]. So, the pulling of the water is to equalize the pressure between the inside and the outside.

The nature of explanatory connections or big ideas made. When constructing CIJ explanation maps, it was evident that, at times, several structural elements together made an explanatory connection or a big idea. Note that not all structural elements produced by participants were included in a connection. Sometimes, participants provided isolated statements that were not necessarily connected with other statements.

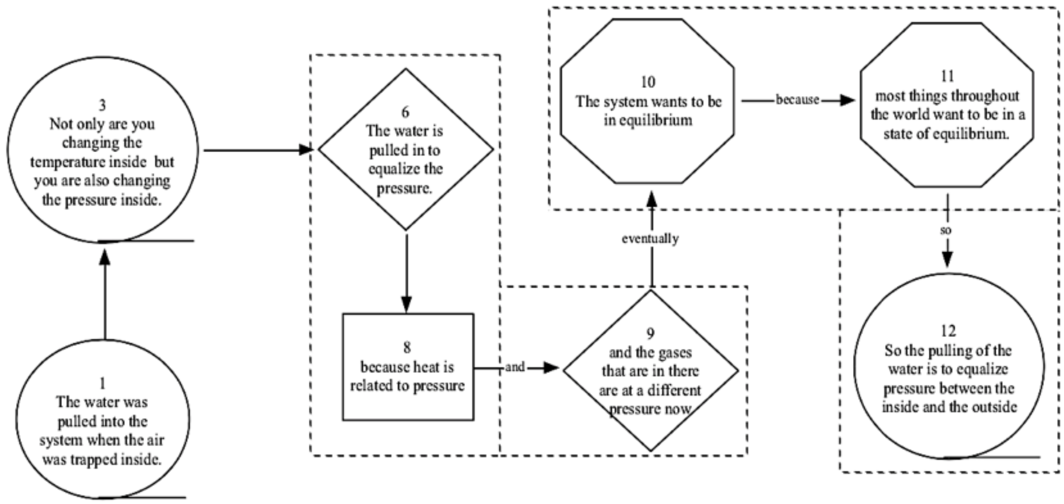


Figure 4.11 Tracy's explanation map. An example of a teacher's map showing teleological and anthropomorphic statements in her scientific explanation.

Recall that in explanation maps, several structural elements included in a dotted rectangular shape refer to these connections (see Table 3.1 for more details). An examination of the structural elements used by all participating groups to make explanatory connections in the CIJ demonstration revealed that scientists did not make more connections than students and teachers. In fact, teachers made more connections (39 connections) than both students (34 connections) and scientists (33 connections). However, scientists' explanatory connections included significantly more structural elements than those made by teachers and students. What is more, scientists' connections included more pieces of knowledge and less inferences and observations; whereas in students' and teachers' explanatory connections, inferences were the major structural elements used.

As evident in Figure 4.12, 36% of the total structural elements produced by scientists to make explanatory connections in the CIJ were pieces of knowledge, followed by laws and lawlike statements (21.5%), inferences (19%), and necessary conditions (14%), with only 5% being observations. On the other hand, teachers' explanatory connections were made of various structural elements in which 46% of these elements inferences, followed by pieces of knowledge (19%), laws and lawlike statements (9%), and 7% were necessary conditions. Finally, students' explanatory connections included various structural elements where 45% of these elements were inferences, followed by teleological/anthropomorphic statements (15%), observations (13%), laws and lawlike statements (11%), and only 9% being pieces of knowledge.

In summary, an analysis of explanation maps in accordance with the NOSE framework indicated that participants tended to use certain structural elements when building their explanations of the CIJ. While students and teachers tended to rely mostly on inferences

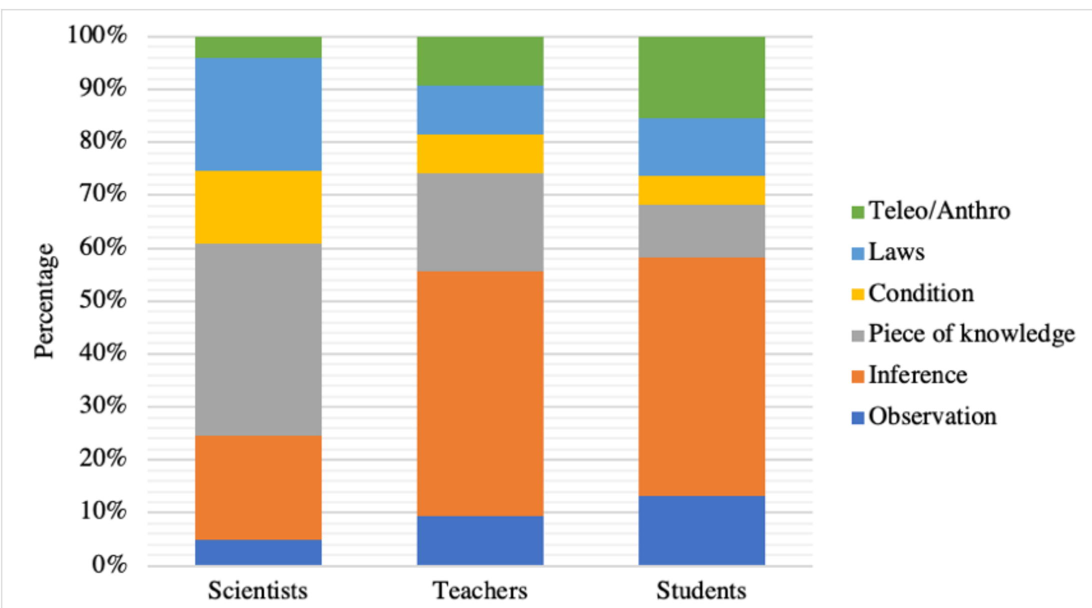


Figure 4.12 Participants' explanatory connections in the CIJ. The structural elements are based on percentage frequencies.

supported by observations, scientists tended to use less observations and inferences, and more pieces of knowledge and lawlike statements – where the two latter structural elements were used to support their fewer observations and inferences. Furthermore, even though scientists produced less inferences and observations, they demonstrated the highest ratio of accuracy and relevance of these elements as compared to those produced by teachers and students. Finally, the use of teleological and anthropomorphic statements in explanations of the CIJ phenomenon was not necessarily an indication of a non-scientific explanation. Two teachers seemed to use such statements as a pedagogical tool where these statements were linked with lawlike statements and pieces of knowledge. In fact, there is an ongoing debate on the validity of teleological explanations in science that requires further examination, for a NOSE analysis perspective.

Basic features of explanation maps within groups. Several major aspects emerged when examining CIJ explanations maps. These were the: (a) presence of the aforementioned structural elements, mainly, the use of observations, inferences, and pieces of knowledge; (b) use of causal links and causal connections; and (c) types and quality of the explanations constructed. The following sections present these results within each participant group.

Practicing scientists' explanations. Table 4.3 shows that a total of 121 statements were observed in scientists' maps; of those 38 (31.4%) were pieces of knowledge, 29 (24%) were inferences, 24 (20%) were laws and lawlike statements, 14 (11.5%) were necessary conditions, 12 (10%) were observations, and only 4 (3.3%) were teleological and anthropomorphic statements. Of the pieces of knowledge, scientist explanation maps showed mostly the use of accurate relevant PK (89.4%) and of lawlike statements, with mostly the use of general deterministic laws and natural regularities (87.5%).

First, it is worth mentioning that when asked, all participating scientists said they had not seen the CIJ before. Nine scientists said that the scientific concepts associated with the demonstrations were familiar to them, while one said that they were not. Sylvia, the participant scientist who was not familiar with the science behind the CIJ, had a background in neurotoxicology and expressed that she was not comfortable with concepts in physics and chemistry. Note that Sylvia was the only participating scientist who used mainly teleological and anthropomorphic statements in her non-scientific explanation of the CIJ phenomenon.

Table 4.3

Frequency of the Main Structural Elements of the Burning Candle Phenomenon Within Groups

	Scientists		**Teachers**		**Students**	
Structural Elements	*f*	%	*f*	%	*f*	%
Observations	12	10	22	17	28	33
Inferences	29	24	54	41	57	44.5
Pieces of Knowledge	38	31.4	25	29	11	8.5
Lawlike Statement	24	20	13	10	9	7
Necessary Condition	14	11.5	9	7	9	8
Teleological/Anthropomorphic Statements	4	3.3	9	7	14	11
Total	**121**		**132**		**128**	

An overall examination of scientists' maps indicated that 9 of the 10 generated explanations were in fact scientific explanations; whereas one was a non-scientific teleological/anthropomorphic explanation (see Table 4.4). Of the remaining nine explanations, two were adequate, three were mostly adequate, and four were partially adequate. In addition, 5 of the 9 explanations included causal connections in the form of causal processes and interactions that lead up to the phenomenon-to-be-explained. Four of these explanations were subsumed under natural laws and lawlike statements, thus producing Causal-Mechanical Deductive-Nomological (CMDN) explanations and one was subsumed under a probabilistic/statistical type law, thus producing a Causal-Mechanical Inductive-Statistical (CMIS) explanation. However, not all of these five explanations were adequate; only 2 of the 5 CMDN and CMIS explanations were adequate, two were mostly adequate, and one was only partially adequate. For instance, recall Sara's explanation map in Figure 4.4. While she used multiple cause-effect relationships to explain why the water rose and supported her explanation with a lawlike statement, Sara still explained the CIJ phenomenon due to the consumption of oxygen that created a vacuum, which in turn caused a pressure difference. Sara's explanation of the CIJ phenomenon was partially adequate, in accordance with the NOSE framework. On the other hand, Stefan used multiple cause-effect relationships subsumed under adequate laws and lawlike statements, thus constructing an adequate CMDN explanation of the CIJ phenomenon. The following excerpt is Stefan's explanation and its corresponding explanation map (Figure 4.13):

> By placing the jar, we limit the amount of oxygen available for the candle. And as it burns through the oxygen then the amount of gas is less, but still within the same volume of the jar. But the water level does not rise instantly but continues to rise after the flame goes out. So, we have a gas that becomes lighter because there is less of it. And according to the equation PV = nRT, the volume is changing but then CO2 is being formed. So, it is not the volume; maybe it is temperature since it is PV= nRT. So, the gas would expand because of the temperature increase. So, as T increases V increases in theory. So, the gas is being heated up by the flame. And the hotter the gas is, the more room it takes. And when the flame disappears there is no source of heat anymore, and the gas is going to

cool down. And by cooling down it is going to retract. So, the pressure is going to be lower inside, and the water goes up to the place of lower pressure in order to reach equilibrium.

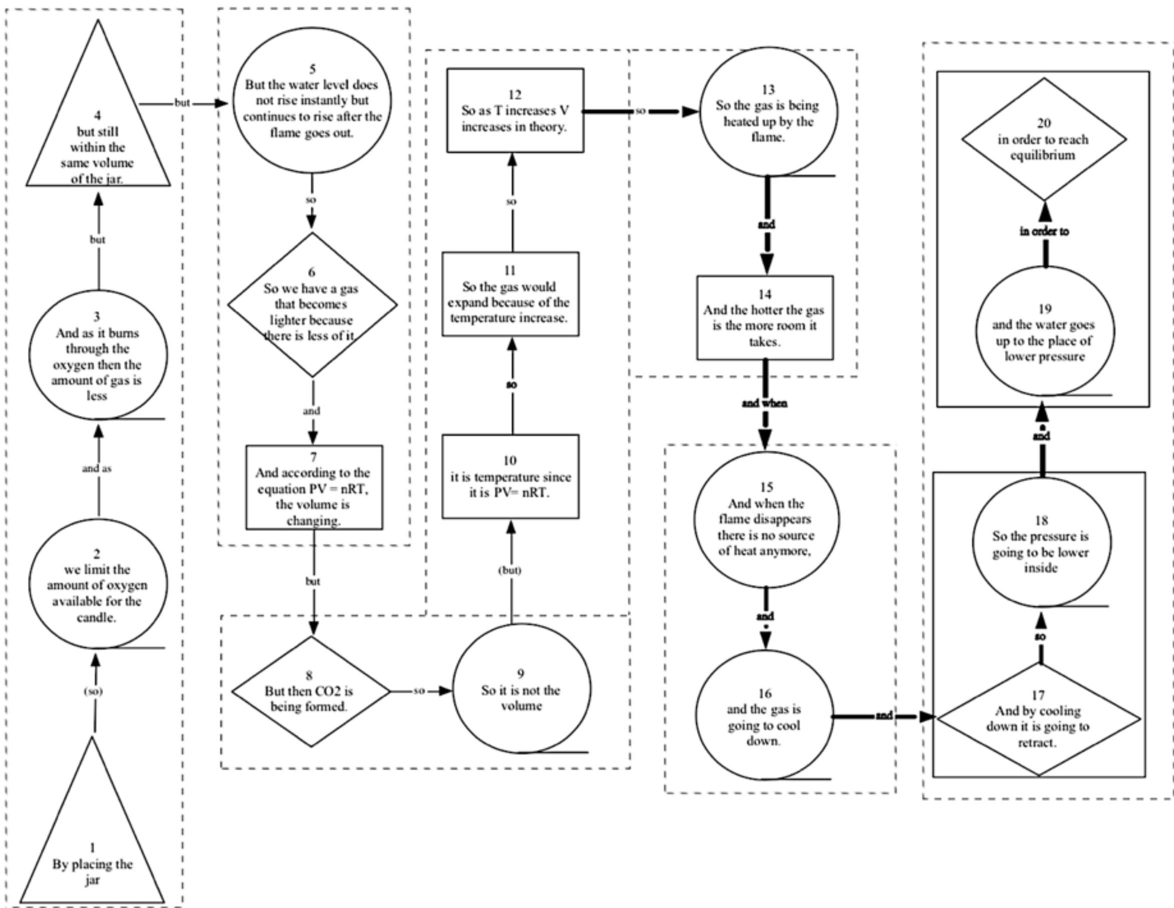


Figure 4.13 Stefan's CIJ explanation map. An example of a scientist's map showing an adequate CMDN explanation of the CIJ phenomenon.

Simple cause-effect relationships were identified in two of the scientists' explanations. Both Stanley and Stella used simple causality supported by general laws to explain the CIJ phenomenon, thus producing Causal Deductive-Nomological (CDN) explanations. However, both of these scientists' explanations were partially adequate. In particular, Stella included several irrelevant statements to explain the CIJ phenomenon mentioning that she was unsure whether the water rose due to oxygen consumption, expansion of the gas, or osmosis. She used simple cause-effect relationship to explain that the water "*caused* [emphasis added] a better seal to the jar, [so] there was no more oxygen able to get in." Stella's full explanation and its corresponding explanation map (Figure 4.14) are presented:

> The flame causes enough heat inside the jar so the air inside of the flask expands and moves up. The water moving in caused a better seal to the jar. There was no more oxygen able to get in it. It might have something to do with osmosis. There is less oxygen, and water moves from low concentration to high concentration. So, after the candle goes out there is condensation and there is enough surface tension that it pulled the water in with it. If we maintain the same temperature of the jar so you can prevent condensing of the gas due to heat change and there would still be the same amount of oxygen in the jar. If the water still goes up, then it is not due to osmosis but due to expansion of the gas.

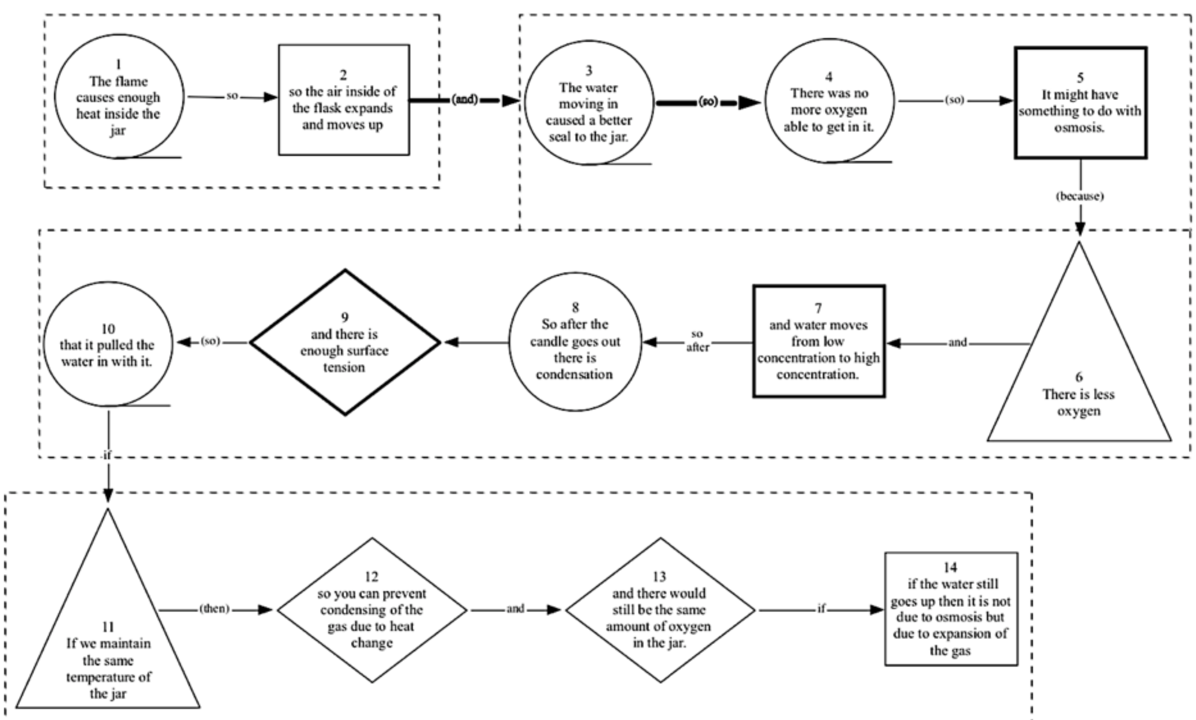


Figure 4.14 Stella's CIJ explanation map. An example of a scientist's map showing a partially adequate CDN explanation of the CIJ phenomenon.

Finally, in the remaining 2 of the 9 scientists' scientific explanations both Sophia and Selena explained the CIJ phenomenon as a consequence of general laws and lawlike statements with no causal connections, thus producing Deductive-Nomological (DN) explanations. In particular, Selena explained the water rising due to temperature change that was related to change in pressure without further accounting for the temperature change: "Inside the jar is now warmer than outside. And the change in temperature is related to change in pressure. I am thinking of the ideal gas law." Sophia, on the other hand, explained the CIJ phenomenon due to difference in pressure and added a lawlike statement related to pressure equilibrium. In accordance with the NOSE framework, while both Selena's and Sophia's explanations were DN explanations, Selena produced a mostly adequate explanation while Sophia's explanation was only partially adequate – pertaining the pressure different and pressure equilibrium.

High school science teachers' explanations. Ten high school science teachers generated CIJ explanations. Table 4.3 shows that a total of 132 statements were observed in teacher maps; of those 54 (41%) were inferences, 25 (29%) were pieces of knowledge, 22 (17%) were observations, 13 (10%) were lawlike statements, 9 (7%) were necessary conditions, and 9 (7%) were teleological or anthropomorphic statements. Of the 54 inferences, teacher explanation maps showed mostly the use of accurate and relevant (72%) inferences, and all lawlike statements in teacher explanation maps were general deterministic laws and natural regularities.

Unlike scientists, teachers' familiarity and prior knowledge with this experiment was evident. In particular, three teachers said that they had seen and done the activity before. Thomas and Trevor were also familiar with the misconceptions associated with it; in their explanations they noted why the water rising phenomenon could not have been the result of the oxygen consumption only. Nonetheless, only Thomas provided an adequate CMDN explanation of the phenomenon; whereas Trevor's was mostly adequate. The remaining eight teachers said that they had seen the CIJ or something similar to it before, but they did not remember where.

Table 4.4

Nature and Quality of the Candle Burning Explanations by Explanation Type Generated by Practicing Scientists (S), High School Science Teachers (S), and Freshman Students (F)

	DN			**Causal**			**CDN**			**CMDN**			**CMIS**			**Total (By Quality)**
	S	**T**	**F**	**S**	**T**	**F**	**S**	**T**	**F**	**S**	**T**	**F**	**S**	**T**	**F**	
Adequate	--	--	--	--	--	--	--	--	--	2	2	--	--	--	--	**4**
Mostly Adequate	1	--	--	--	1	--	--	--	--	1	1	--	1	--	--	**5**
Partially Adequate	1	1	3	--	--	2	2	2		1	2	--	--	--	--	**14**
Inadequate	--	--	1	--	--	--	--	--	1	--	--	--	--	--	--	**2**
Total (By Type)		**7**			**3**			**5**			**9**			**1**		

*Scientists generated nine scientific explanations, teachers generated nine scientific explanations, and students generated seven scientific explanations. Hence, a total of 25 out of the 30 participant-generated explanations were scientific explanations.

Some said they might have watched it on YouTube, while others said they might have seen it in some professional development workshop but said that they did not remember what happened or why it happened. Trevor adequately explained that the consumption of oxygen did not lead to the rising of water. However, he related the pressure change to the water phase change from gas to liquid (due to temperature change):

> As the flame is burning, you are heating up the air. But hot air has more pressure. So, if you notice on the inside of the jar there is some humidity that has condensed. Because when the gas condenses there is a big pressure change from going from a gas to a liquid. And that would cause the lower pressure and that's why the water goes inside.

Hence, an overall examination of teachers' maps indicated that 9 of the 10 generated explanations were in fact scientific explanations; whereas one was a non-scientific teleological explanation. As evident in Table 4.4, 2 of the 9 explanations were adequate, two were mostly adequate, and five were only partially adequate. Similar to scientists, 5 of the 9 explanations included causal connections in the form of causal processes and interactions that lead up to the phenomenon-to-be-explained subsumed under natural laws and lawlike statements, thus producing CMDN explanations. However, not all of these five explanations were adequate: only 2 of the 5 CMDN were adequate, one was mostly adequate, and two were only partially adequate. For instance, Tucker constructed a partially adequate explanation where he explained the CIJ phenomenon due to oxygen consumption instead of temperature change. He then explained pressure differential and subsumed the phenomenon to laws of entropy and thermodynamics. Tucker's full explanation is presented (see Figure 4.1 for its corresponding explanation map):

> The system was not closed because water was able to come in which means gas can come in and out. There was also a bit of an evaporative effect because we can see condensation of the water to gas on the jar. And the pressure difference is directed inwards which causes the water to move inside. Pressure is caused from the reducing of oxygen within the jar. Inside the jar there was less pressure, which as a result caused to flood inwards and now it is at an equal point between the inside and the outside of the jar. The nature of

> entropy tells us that we are looking for the most equivalent state or the state of equilibrium between the two and that is where the pressure gradient on the inside matches the pressure gradient on the outside. And that follows the laws of thermodynamics.

Simple cause-effect relationships were observed in three of the teachers' explanations. Both Tina and Tyson used simple causality supported by general laws to explain the CIJ phenomenon, thus producing Causal Deductive-Nomological (CDN) explanations. On the other hand, Tanya produced a mostly adequate causal explanation where she included mainly inferences and pieces of knowledge connected by causal links. Both of Tina's and Tyson's CDN explanations were only partially adequate. Like a considerable number of other participants, Tina explained the CIJ phenomenon due to the consumption of oxygen only. She supported her explanation by a lawlike statement related the behavior of fluids: "Fluids which are gases or liquids are going to take the path of least resistance, so this is why the water goes to a place with less pressure." On the other hand, Tyson explained that the water rising was due to heat, citing a lawlike statement that "warm air does rise… I know that's for sure." However, he did not further provide any other explanation. The following is Tyson's CIJ explanation with its corresponding explanation map (Figure 4.15):

> So, the heat does suck the water in. Warm air does rise. And there might be enough of a force because of that that's going to suck it in. And so that's an effect of it. You just have that force that is created because fire is a force. Heat is a form of energy. And heat is there, and energy is allowed to suck it up. The one remaining scientific explanation was constructed by Tracy (already presented in Figure 4.11) in which she constructed a partially adequate DN explanation. Tracy explained the CIJ phenomenon citing the

relationship between heat and pressure, but she did not give further explanation to why the water rose. Finally, Tarra, a novice high school science teacher, had started teaching

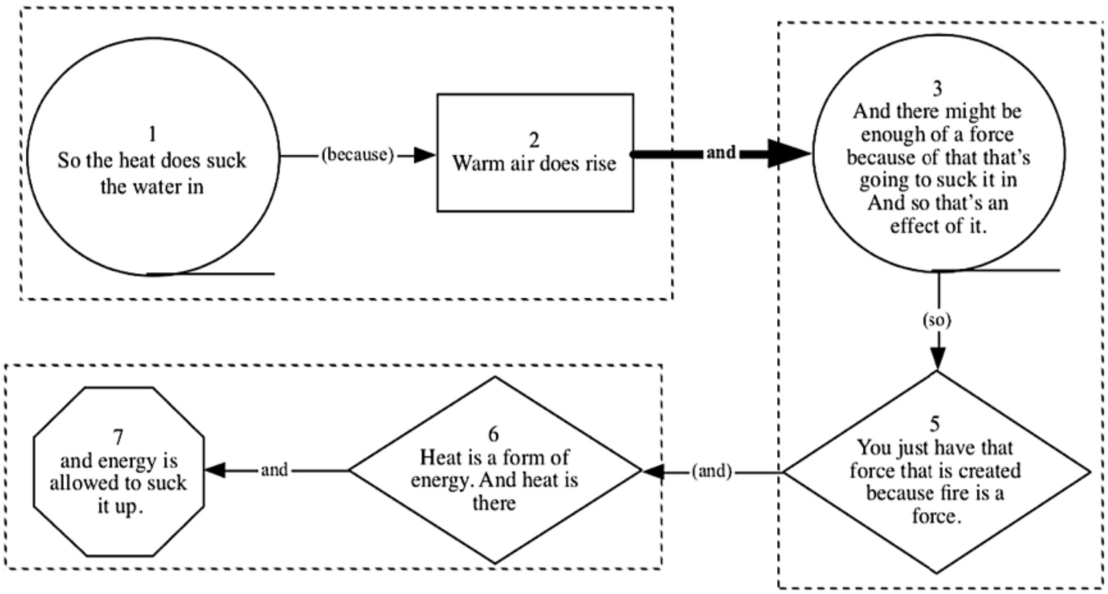


Figure 4.15 Tyson's explanation map. An example of a teacher's map showing a partially adequate CDN explanation of the CIJ phenomenon.

only a few months before participating in this study produced a non-scientific CIJ explanation. Tarra earned an undergraduate degree in Animal Sciences and a master's degree in education. In addition to her job as a high school science teacher, she also worked at a veterinary clinic. She stated that she had seen the CIJ before but that she did not remember where, nor did she remember what happened or why it happened. Tara's explanation was the only teacher-generated non-scientific explanation in which she explained the water rising using irrelevant and inaccurate pieces of knowledge along with anthropomorphic statements that included human-like nature of water molecules (Figure 4.16):

> You put the jar on top of the candle which was lit at the time. And the water was surrounding it in a dish and then slowly the water started getting sucked up into the jar. So now the water is literally inside the jar. So, as the water is coming in it is combining

with other gases within that area. Therefore, it cannot be used in that combustion reaction which can no longer happen because the reactants are limited. So, basically you have two different phases of matter –gas in the jar and it is combined with water – another phase of matter. The water is more dense than the gas therefore there is this relative suction between the two. So, the gas wants to float but there is still other gas that is surrounding it in this outside. So, the water does not really know where to go. And then as the combustion reaction is happening, there are more molecules that are moving a lot faster because they are happier because they are in an environment where they can go wherever they want to go outside. Whereas the gas molecules inside are trapped. So, the water is moving closer to those molecules because it feels that source of pressure.

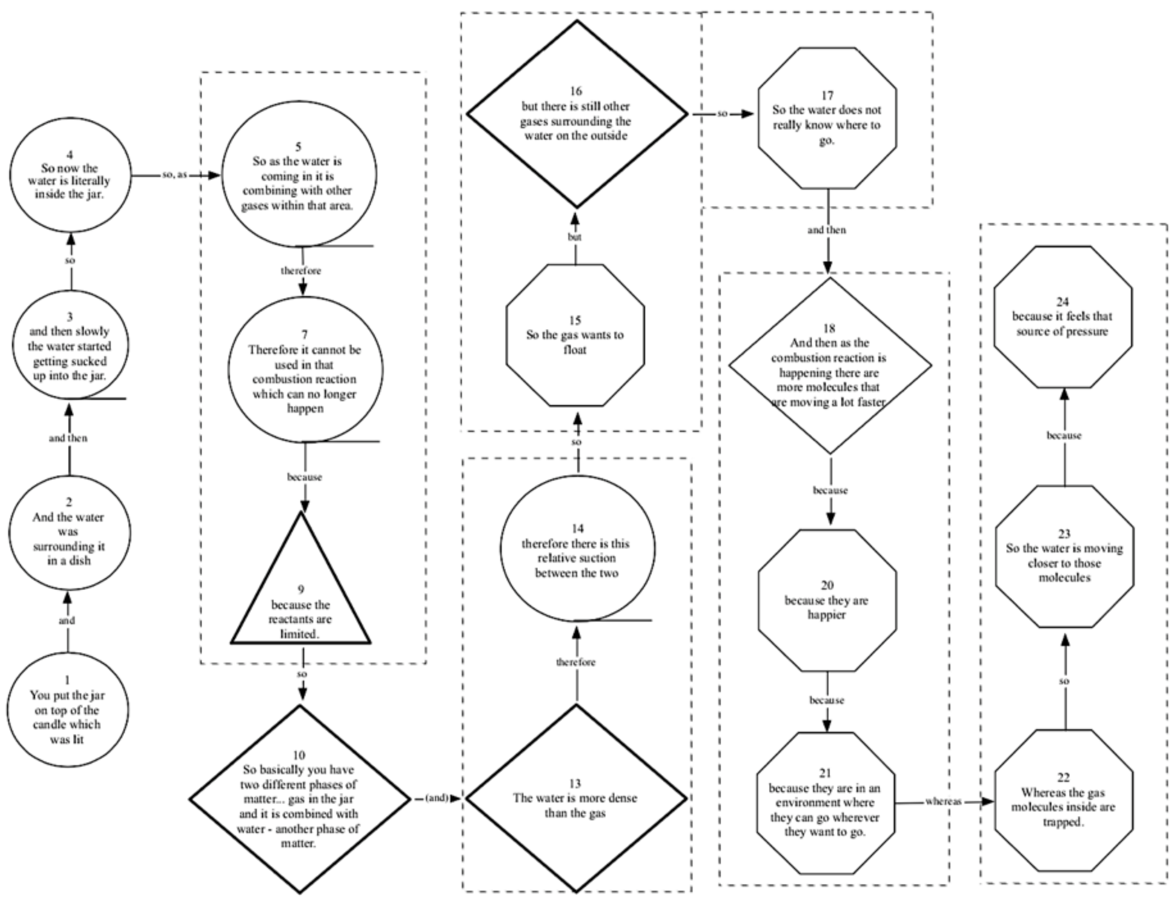


Figure 4.16 Tarra's explanation map. An example of a teacher's map showing a non-scientific explanation of the CIJ phenomenon.

College freshman students' explanations. Ten college freshman students produced explanations of the CIJ. Table 4.3 shows that a total of 128 statements were observed in student maps; of those 57 (44.5%) were inferences, 28 were observations (22%), 14 (11%) were teleological or anthropomorphic statements, 11 (9%) were pieces of knowledge, 9 (7%) were necessary conditions, and 9 (7%) were laws or lawlike statements. Of the 57 inferences, student explanation maps showed that 30 (53%) of them were accurate and relevant, while the remaining 27 (47%) where inaccurate or irrelevant inferences.

When asked, all participating freshman students said that they had not seen the CIJ before. However, one freshman student, Finn, said that it reminded him of another activity: the egg in a bottle. Finn recalled that he did not understand that demonstration when he saw it done. Another student, Farrah, stated that the rising of water due to heat reminded her of the hot air balloon. Farrah also mentioned that she had been in one of those hot air balloons before and had someone explain to her that hot balloon went up because of the heat. She further stated that she was not sure how that happened, however. When asked if they had seen the CIJ or if it was familiar to them, all freshman students noted that they had seen birthday candles before, and they knew why candles blew out.

An overall examination of students' explanations shows that they produced the largest number of non-scientific explanations and inadequate explanations as compared to the other two groups. Furthermore, they used simple cause-effect relationship in their causal explanations. In particular, freshman students generated three teleological explanations that did not scientifically explain the water rising phenomenon. Instead, these three students explained the water rising because "oxygen wanted to reach the flame," or because "the flame needed oxygen, so it pulled the water to it." In addition, two freshman students generated inadequate scientific explanations (one was a DN explanation and the other was a CDN explanation). The remaining five were all partially adequate explanations (four were DN explanations and the other was a Causal explanation).

As mentioned earlier, three freshman students provided non-scientific CIJ explanations. In her explanation, Felicia included many inaccurate inferences, teleological and anthropomorphic statements, and non-scientific causal links. Felicia's CIJ explanation and its corresponding map are presented (Figure 4.17)

> In the first jar that we used I noticed there were bubbles that caused the flame to put out, but at the same time I also saw the water rise in a fast motion. But once it got to the level where there was no more water surrounding the glass, that was what caused the fog that formed. This has to do with concealment of the space. Right away the water was trying to rise up on all of the three jars. At first my initial thought was it was trying to get to the flame but then once you did the other two [jars]. Once the bubbles started, the flame was put out. The water is trying to rise because of the fog. So, the bubbles is what caused the flame to be put out. But then once the water reached the level that's when the smoke and the vapor came out. So, it has to do with something with the vapor and the water. There is a connection. But then once the water reached the level, they are still trying to escape and at the same time they are concealed so they are not being able to escape.

Additionally, 4 of the 7 student-generated scientific explanations were DN explanations in which students explained the CIJ phenomenon as a consequence of a law or lawlike statement without explicitly using any causal links or causal connections. As discussed earlier, none of these explanations were adequate: three were only partially adequate and one was inadequate. Furthermore, Franco constructed the only CDN explanation; however, it was an inadequate one.

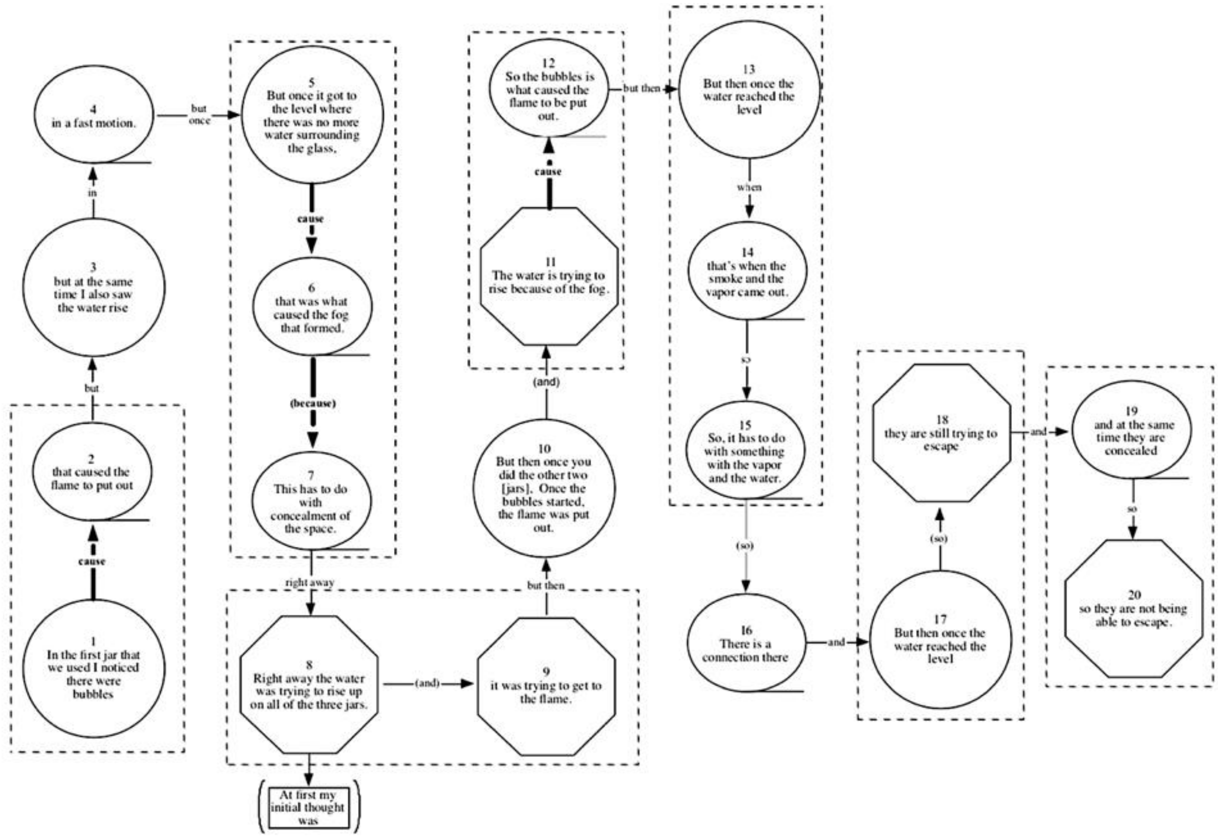


Figure 4.17 Felicia's explanation map. An example of a student's map showing a non-scientific explanation of the CIJ phenomenon.

Franco explained the water rising causally due to water gradient – clearly an irrelevant premise. What is more, he exhibited a clear misunderstanding of the concept of water gradient. In particular, Franco stated that (see also Figure 4.18):

> Water moves on a gradient from high concentration to low concentration. So there was more water outside the jar than inside the jar. Once you put the jar over the candle, you already had a little bit of water trapped inside, there was all the oxygen molecules on top of the water. Once the fire started burning through those oxygen molecules, the atmosphere within the jar, which was in a sense pushing down on the water was removed.

And since water on that gradient, it moves from high concentration to low concentration, it started moving up because now it had more space to move up.

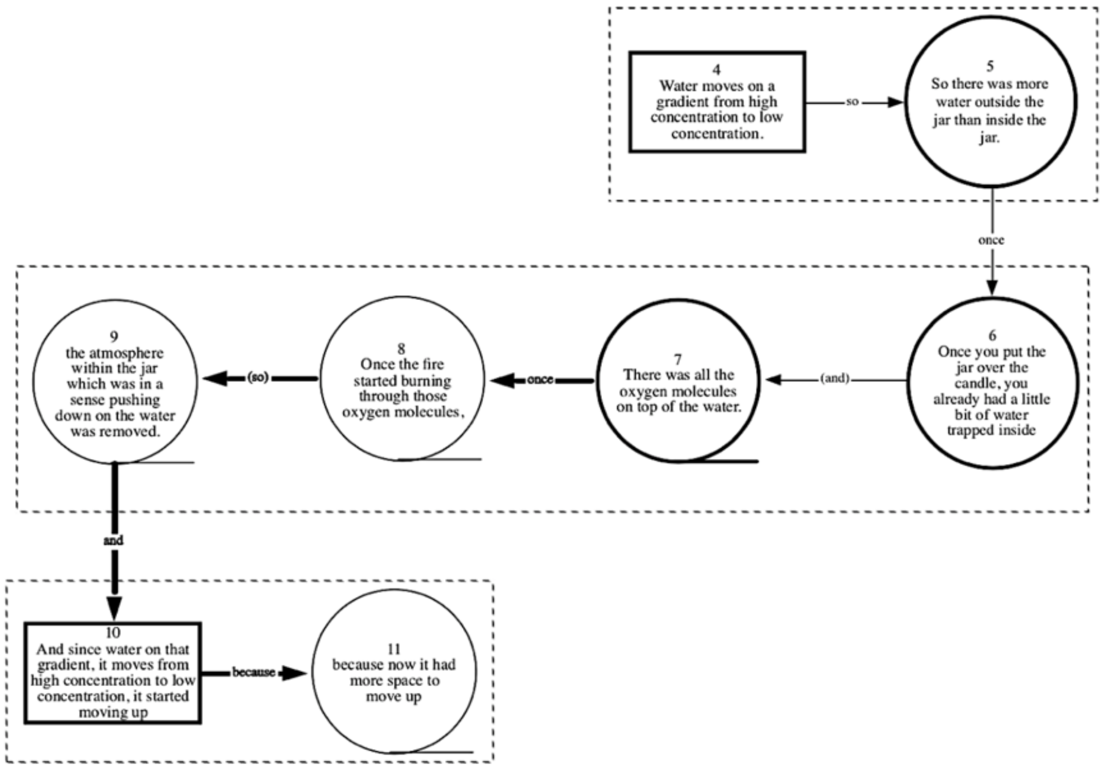


Figure 4.18 Franco's explanation map. An example of a student's map showing an inadequate CDN explanation of the CIJ phenomenon. Bolded shapes refer to irrelevant structural elements.

Finally, Faith and Flynn provided partially adequate causal explanation in which inferences, observations, and necessary conditions were causally linked but were not subsumed under any general laws or lawlike statements. For instance, Faith explained that "when you put the jar over the candle, it creates a vacuum, so it sucks the water… it takes away its source of oxygen because it will suck all of it that is in the jar"; while Fidel stated that "as it [the candle] burns through the oxygen the water gets sucked up because it created a vacuum.

Comparison of the Three Groups. Of the 25 participant-generated scientific explanations of the water rising phenomenon only four were adequate, two of which were provided by scientists and two by teachers. As mentioned earlier, unlike scientists, teachers were familiar with the CIJ and some had previously seen and done the demonstration. In addition, freshman students were the only group who did not produce any CM explanations; instead, they used only simple cause-effect relationships as compared to teachers and scientists. Additionally, unlike teachers and students, all scientists' scientific explanation included pieces of knowledge and laws and lawlike statements.

An examination of all participants' types of explanations shows that CMDN explanation type was the most generated among the three groups (9 out of the 25 explanations were CMDN), followed by seven DN explanations, five CDN explanations, three causal explanations, and one CMIS explanation. Hence, the use of causal links and causal connections (simple and multiple connections) was evident as 18 out of the 25 participant-generated explanations were causal. As would be expected, scientists demonstrated the highest use of multiple causal connections subsumed under laws and lawlike statements, while teachers and students used more simple cause-effect relationships to indicate causality.

Overall, inferences and pieces of knowledge were the major structural elements observed in participant maps with differences in the frequency of each per group, their relevance and accuracy. In particular, scientists tended to use significantly more pieces of knowledge over inferences and observations compared to students and teachers. In addition, scientists tended to support their inferences with lawlike statements as opposed to students who based their inferences on their observations. Additionally, scientists and teachers relied more on prior science content knowledge than students, whereas students used sensory observations and descriptions.

The Dancing Raisins (DR) Scenario

Basic features of explanation maps across groups. In the DR scenario, results from participants' explanations of the water rising phenomenon revealed common patterns across all participants in relation to the use of observations and inferences versus pieces of knowledge, predictions, the use of necessary conditions, and the nature of explanatory connections or big ideas made. Each of these aspects is discussed below.

Observations and inferences. During the first round of interviews participants were asked to provide an explanation in which they described what happened when the researcher added a few raisins to a cup of clear soda (7Up or Sprite) and explained why it happened. Table 4.5 shows that, similar to findings from the CIJ, while constructing DR explanations, students and teachers demonstrated the highest use of observations and inferences compared to scientists: out of a total of 59 observations generated by all participants, 49% (29 observations) were produced by students and 29% (17 observations) by teachers as compared to 22% (13 observations) produced by scientists. What is more, out of 238 total inferences generated by all participants, 40% (96 inferences) were produced by teachers and 33% (77 inferences) by students as compared to 27% (65 inferences) produced by scientists.

Aligned with the NOSE framework, observations and inferences were assessed based on their relevance and accuracy. Results showed that all observation statements generated by participants from the three groups in relation to the DR scenario were accurate and relevant. On the other hand, three types of inference statements were observed among all three groups: accurate relevant inferences, inaccurate relevant inferences, and accurate irrelevant inferences. Interestingly enough, the vast majority of participants' – this time – accurate and relevant observations were related to the bubbles observed during the DR demonstration (recall that many of these same participants did not observe the bubbles in the CIJ). In fact, all 30 participants generated observation statements related to the bubbles in the soda.

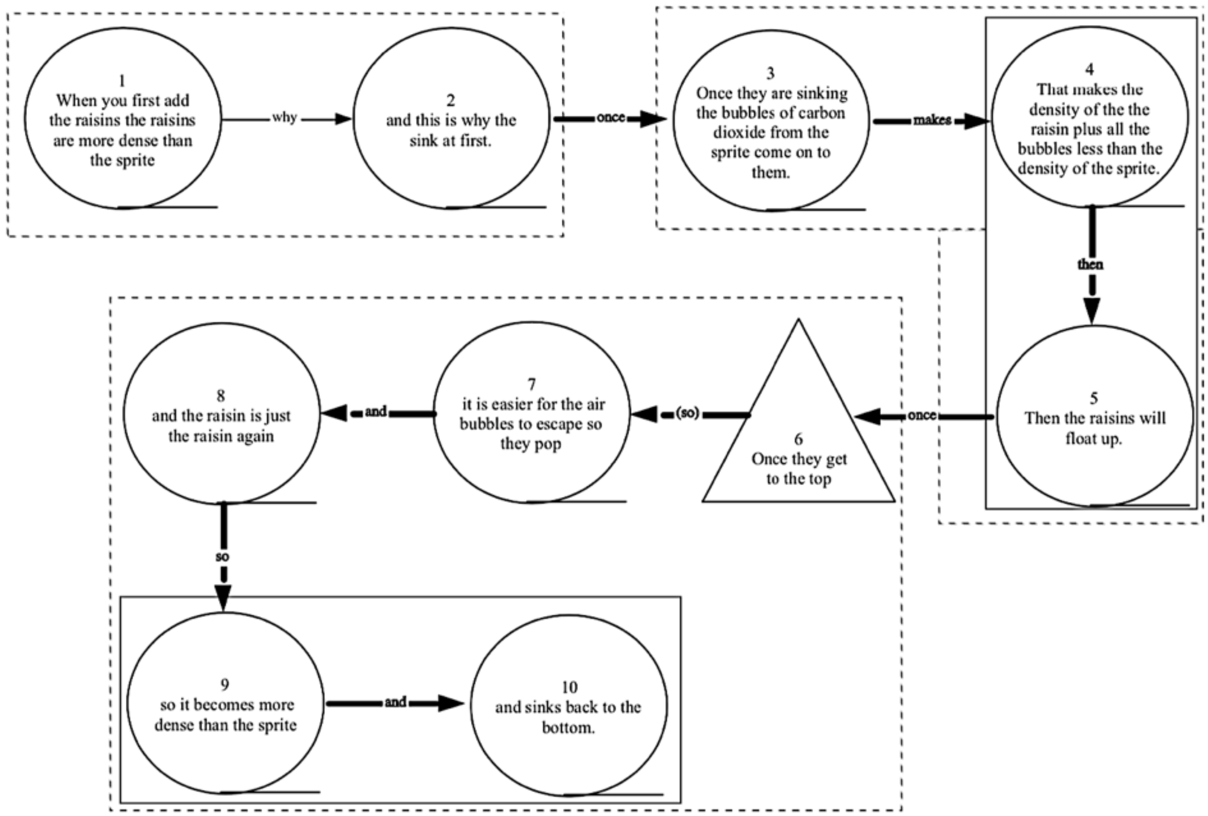


Figure 4.19 Sample Explanation Map of the Dancing Raisins Phenomenon. Explanation is provided by Flynn, a participant freshman student. All shapes contain statements directly excerpted from participant transcripts. Numbers are used to identify the sequence of statements. Different shapes indicate different structural element as can be seen in Table 3.1. Arrows show the directional flow of statements for a participant. Bolded arrows indicate causal relationships. Dotted rectangles containing several structural elements refer to an explanatory connection or a complete idea.

The majority of participants stated that bubbles “adhere”, “attach”, or “stick” to the raisins. For instance, Finn stated that “the bubbles formed on them [the raisins],” Sam said: “When the bubbles are stuck to the raisins, the bubbles and the raisins move together,” and Thomas said that “you can see that more bubbles will form on the raisins.” Some participants, mainly scientists, further stated observations related to the number of bubbles on the raisins, the behavior, size, shape, and position of the raisins, and the coating of the surface of the raisins.

These observations were often associated with various pieces of knowledge and inferences made by participants to explain the phenomenon - and this is elaborated later in this chapter.

Table 4.5

Frequency of Occurrence of Major and Minor Structural Elements Used by Participants in Explanation Maps of the Dancing Raisins Scenario Across the Three Groups

	Scientists	**Teachers**	**Freshman Students**	**Total**
Structural Elements				
Observation	13	17	29	**59**
Inference	65	96	77	**238**
Piece of Knowledge	41	24	20	**85**
Necessary Condition	17	9	11	**37**
Law and lawlike statement	29	22	18	**69**
Prediction	9	6	7	**22**

An interesting finding reveals that while all participants noted that the raisins first fell to the bottom, then went up to the top and then fell back to the bottom, participants form the three groups still did not make the same kinds of observations. While students stated that *all* the raisins behaved in the aforementioned way, scientists and teachers specified that *some* raisins exhibited an up and down movement while others did not. Scientists and teachers, after stating such observations, provided further explanations to why some but not all raisins behaved that way. For example, Tarra stated:

> After you put the soda in and then had the raisins in there, most of the raisins went to the bottom right away and then they have little bubbles on them. And some of them are randomly coming back up to the top and then coming back down again. This one over here has been doing this a lot more than the others.

Similarly, Saul stated:

> Some of them [the raisins] do float but some of them stay at the bottom while some jump up and down. For those that are smaller and lighter, the bubbles can keep them at the top while the ones that are heavier are too big to stay at the top so they go back to the bottom.

It is worth mentioning that the raisins behaved pretty much the same way when the demonstration was performed with all 30 participants. In other words, an accurate observation of the behavior of the raisins that reflected what actually happened would be that some and not all of the raisins behaved a certain way.

As evident in Figure 4.20, the vast majority of inferences generated by all three participant groups were relevant to the phenomenon-to-be-explained: approximately 98% (234 inferences) were relevant to the DR phenomenon and only 2% (4 inferences) were irrelevant. Of the 234 relevant inferences, 82% (191 inferences) were accurate while 18% (43 inferences) were inaccurate. Recall that scientists produced the least number of inferences among the three participating groups. Additionally, scientists generated the highest ratio of accurate inferences: 85% (55 out of 65) of scientists' inferences were accurate and relevant, as compared to 82% (80 out of 96) of teachers' inferences and 73% (56 out of 77) of students' inferences. While the majority of inferences produced by all participants were accurate and relevant, students and teachers produced significantly more inaccurate inferences than scientists. In particular, 19 inaccurate inferences (44%) of all inaccurate relevant inferences were produced by students and 15 (35%) were produced by teachers, as compared to 9 inaccurate inferences (21%) produced by scientists. Finally, a total of four accurate but irrelevant inferences were generated; two of which were produced by a student, one by a teacher and another by a scientist. Franco, Sylvia and Tina were the three participants who generated irrelevant inferences while generating their DR explanations. Franco stated that the bubbles in the soda were oxygen bubbles, and so “when you have oxygen you put it under water it floats back to the top because oxygen is lighter than water. So oxygen is lighter than Sprite,” Sylvia explained the initial sinking of the raisins “because you are dropping them from a higher point than the 7UP,”and Tina described the movement of the raisins as “almost using a convection cycle going up and down.”

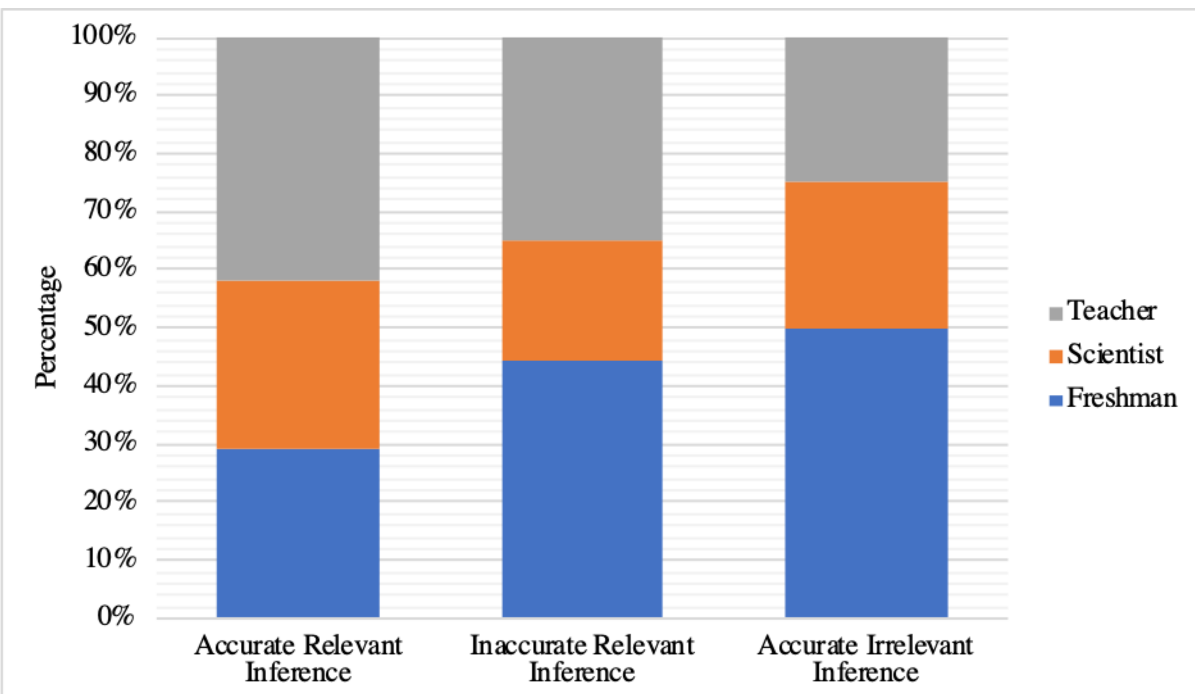


Figure 4.20 Participants' types of inferences. The types are based on percentage frequencies.

Another finding revealed that, in general, inferences constructed in the DR scenario included common ideas related to the DR demonstration across the three groups. In particular, a vast majority of participants' inferences in this regard included inferences about density. Other common, albeit less frequently observed, inferences were related to mass, weight, and gravity, heaviness and lightness, and buoyancy. In fact, 7 of the 10 students, 7 of the 10 teachers, and 9 of the 10 scientists explicitly mentioned density while constructing their DR explanations. In particular, after observing the raisins initially sink, the majority of participants made density comparisons between the raisins and the soda. These inferences were not necessarily accurate and relevant. However, scientists' density-related inferences tended to be more accurate than those generated by students and teachers. For instance, Faith stated that "all the raisins will sink because they are denser or heavier than the 7UP." However, she later added that "they [the raisins] are just not buoyant, so they sink" instead of making a comparison between the buoyant force and the weight of the raisins. Franco, another participant student, first explained that "they [the raisins] first sink because of the whole density of the raisin thing"; however, he then made inaccurate inferences such as: "the raisin is *large* [emphasis added] enough to sink to the bottom," and "the raisins go back to the bottom because they are *heavy* [emphasis added]." A third student, Filip, explained the DR phenomenon by citing Archimedes principle, but then later explained the sinking due to the weight of the raisins. He also added other inaccurate inferences related to a constant frequency with which the raisins were *oscillating*, and the nature of the bubbles that got stuck on the raisins. Filip's complete explanation and its corresponding explanation map are presented (see Figure 4.21):

> The raisins first sink to the bottom because of the weight of the raisins. Gravity is pulling downwards. The weight is mass of the raisins into gravitation, so they are going down

because of their weight. The raisins are going up and down; they are oscillating at some frequency which might be constant. The raisins are going up because the air bubbles are getting stuck to them. The air bubbles don't have that much mass. Recalling Archimedes principle where the density is less and the volume is large. So here less mass is getting distributed in a larger volume. So because of that we get a large force that is pushing it upwards. It is like when we wear life jackets and jump into the ocean. The life jackets keep you afloat because they have a lot of air in there. So the air bubbles here are trying to act as life jackets for these raisins.

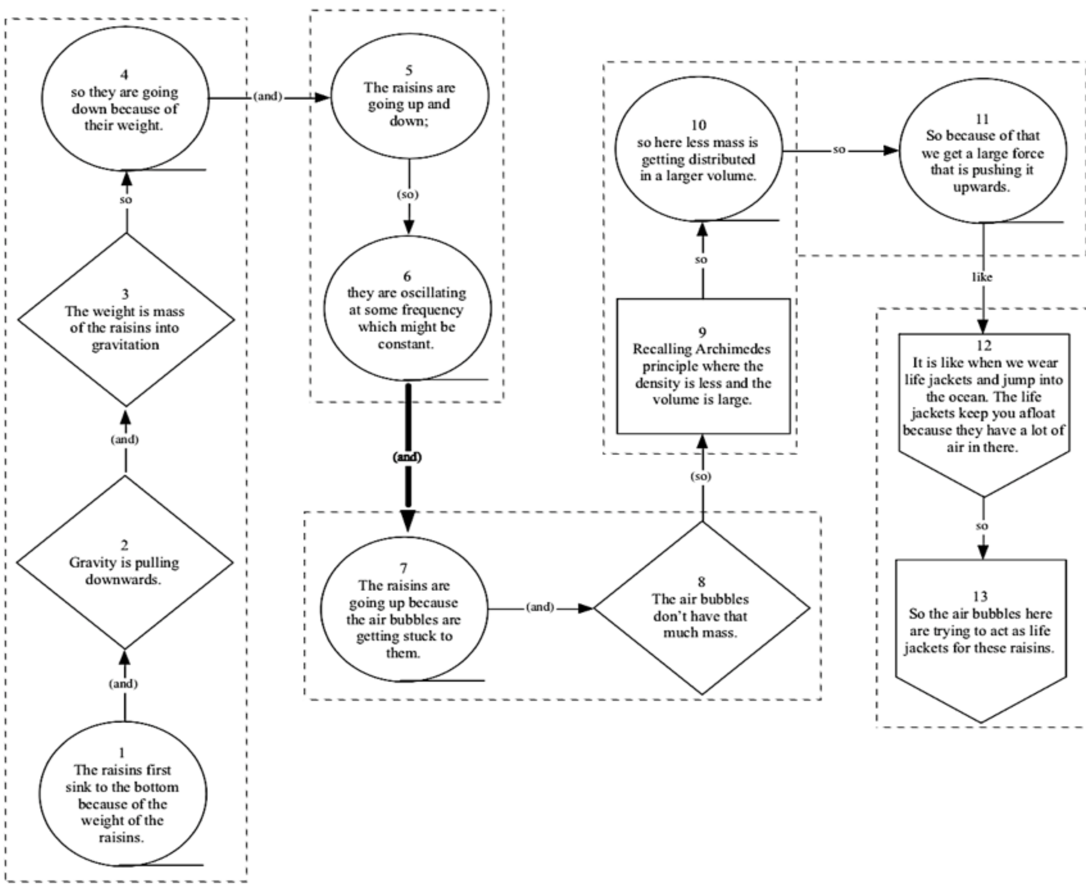


Figure 4.21 Filip's explanation map. An example of a student's map of the DR demonstration.

Other students made inferences pertaining the weight, mass and/or gravity of the raisins while explaining their initial sinking. Felicia stated that "they [the raisins] go down right away because of the mass"; while Farrah said that "the majority went down because of their weight, but some of them did float because they were *lighter* [emphasis added]." These inferences were inaccurate as the weight alone does not determine if an object would sink or float. In fact, objects that sink or float have gravitational force pulling them downwards. What determines floating or sinking is the relationship between the gravitational pull on the object and the upward force of the liquid it displaces.

Similar to students, a number of teachers generated inaccurate inferences related to density, mass, weight, and gravity, heaviness and lightness, and buoyancy. For instance, even though Tucker explained that the DR phenomenon "has to do with density," he included a number of irrelevant and inaccurate statements that rendered his DR explanation inadequate. In particular, Tucker mentioned that there was a change in the density of the raisins, rather than a change in the overall density of the system: raisins-bubbles. Furthermore, he explained this change due to the raisins membrane exchange. He made inferences related to the water molecules being able to go in and out of the raisins in addition to the bubbles surrounding the raisins – that together caused the density to change. Finally, he cited principles of adhesion and cohesion that also lead to change in density of the raisin. Tucker's full explanation and its corresponding explanation map are presented (see Figure 4.22)

> The raisins initially sink to the bottom and then they rose to the top and now they continue to go up and down. They continue to oscillate in the sprite. This has to do with density. You need to see the membrane exchange which leads to a change in the density of the raisins. Because initially it sinks to the bottom, so that tells me that the raisin is

> more dense than sprite. But the carbonation and the molecules within the soda are interacting within the raisins. So there seems to be an exchange where initially it goes in and this leads to a decrease in density, which makes the raisins rise. But when they rise, they then come back down which tells me that their density increases back again. The water molecules should be able to go into and out of the raisin. And you can see bubbles surrounding the raisin. But at the same time it could also be the carbon dioxide that is in the solution. It could be traveling in and out of the raisins. What's surrounding the raisins are water molecules that are interacting with CO2 as well. CO2 should be able to make it across the membrane of the raisin. When enough bubbles escape the raisins, the raisins go to the bottom. So that tells me that the bubbles are causing the raisins to go up. And as more bubbles surround the raisin, you are seeing principles of cohesion and adhesion forming this hydration shell that is leading to decrease in density of the raisin. So then it floats.

Second, unlike scientists, students and teachers tended to make inferences to qualify Tina, another participant teacher, generated a mix of accurate, inaccurate and irrelevant inferences while explaining the DR phenomenon. While Tina produced density-related inferences, like Tucker, she said that "the density of the raisins is changing," and that "the content of the raisins is changing" instead of considering the density of the overall bubble-raisin system. Tina further made inaccurate inferences related to the behavior of the raisins, describing it as a "convection cycle" and that "the gas contained in the carbonated beverage looks like is diffusing."

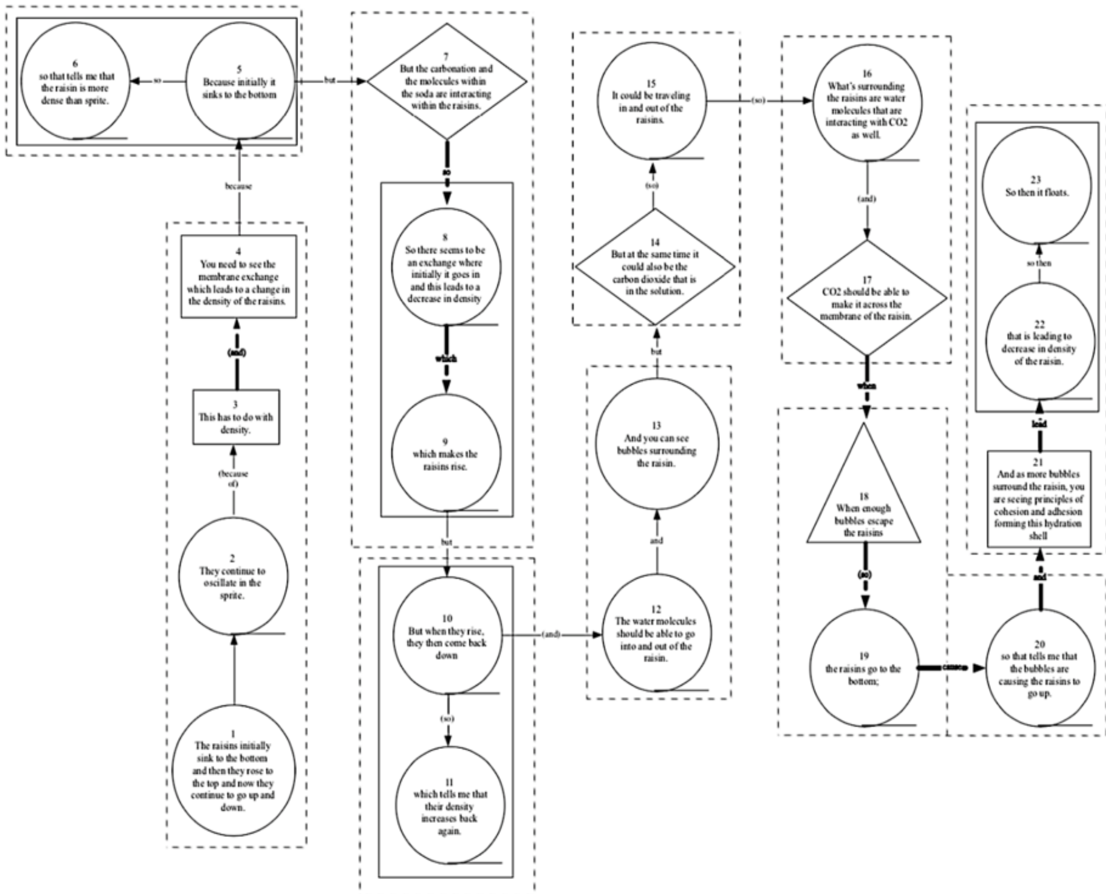


Figure 4.22 Tucker's explanation map. An example of a teacher's map of the DR demonstration.

Tina's full explanation and its corresponding explanation map are presented (see Figure 4.23):

> They first sink then they float and then sink back to the bottom. They are almost using a convection cycle going up and down. They are filling with gas and then losing the gas at the top. As they reach the surface of the water, they release the gas and then they sink back. And when they sink they gain more gas. The gas contained in the carbonated beverage looks like is diffusing. The raisins are originally more dense than the sprite, but then as more gas diffuses into them they lose their density and rise to the surface. But then once they are at the surface, the gas can leave the raisin, and they then become dense again and sink until they diffuse more gas and they go back up. The density of the raisins

is changing as the gas diffuse in it. The content of the raisins is changing otherwise why would it float? It is floating because it is less dense than the sprite because it is now filled with gas.

Other participant teachers made inference statements about the sinking and floating of the raisins in relation to how heavy or light they were in addition to how dense they were. For example, Tyson started his explanation by saying that "the raisins start off being denser," but then later added inference statements such as: "and then the bubbles come off when they hit the surface, so they are heavier, so they come back down." Note that past research has shown that people commonly believe that heavy objects sink and light objects float regardless of their size, shape or the type of material used to make them (e.g. Biddulph & Osborne; 1984; Mitchell & Keast; 2004). Hence, analysis of participants' explanations in this study focused on these misconceptions and how they affected the quality and nature of the explanations they provided.

Scientists were the only participant group that tended to use density and buoyancy-related inferences instead of making inferences about mass or weight in relation to sinking and floating. In fact, only 1 out of the 9 scientists, Sylvia, did not explicitly use density-related inferences to explain the behavior of the raisins. However, she used both size and mass properties: "For those that are smaller and lighter, the bubbles can keep them at the top; while the ones that are heavier are too big to stay at the top so they go back to the bottom." Sylvia's explanation was only partially adequate because she also explained the initial sinking of the raisins due to the fact that they are "heavier than the 7UP' and "because you are dropping them from a higher point than the 7UP" that "initially brings them to the bottom of the glass" – where the latter two inferences were clearly irrelevant. Sylvia's complete explanation and its corresponding explanation map are presented (see Figure 4.24).

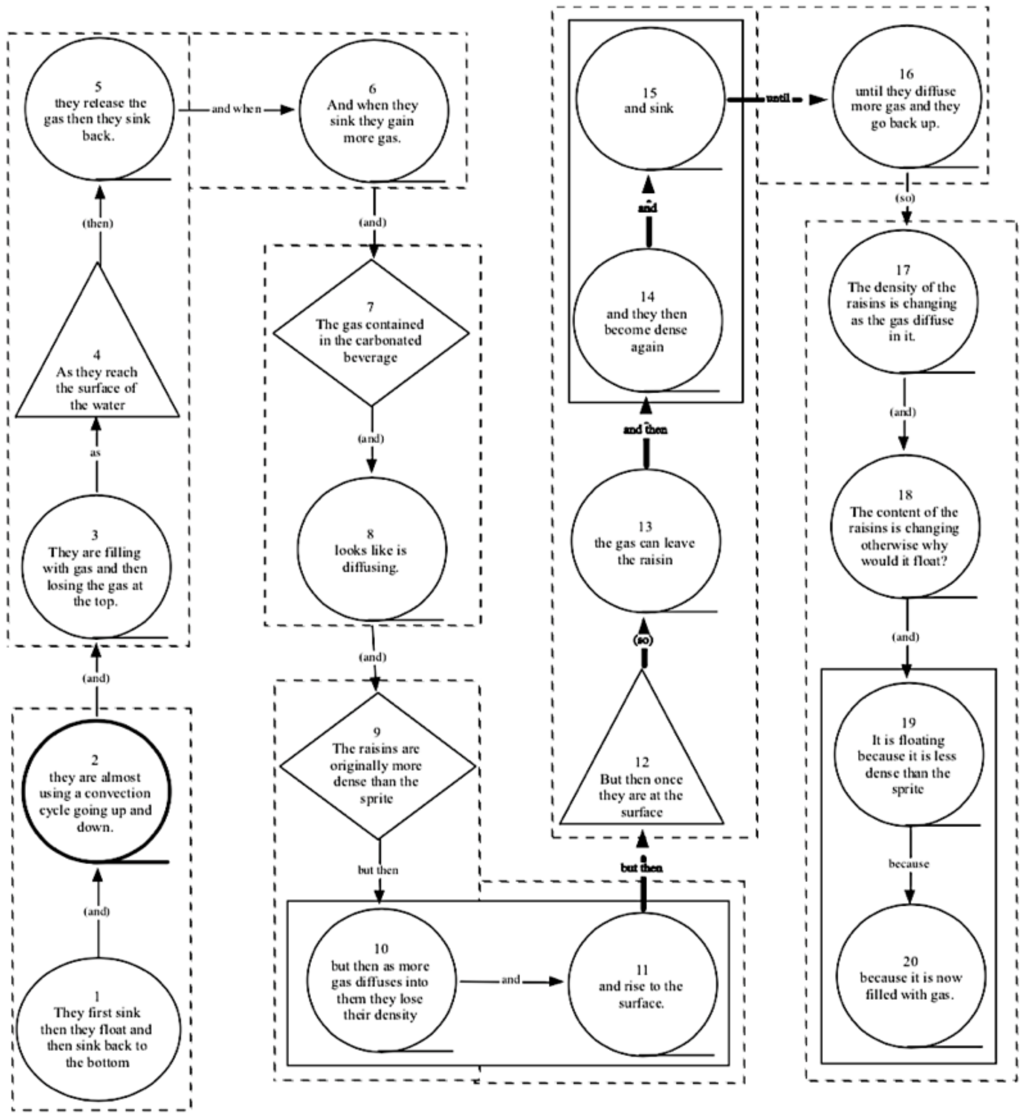


Figure 4.23 Tina's explanation map. An example of a teacher's map of the DR demonstration.

When you put the raisins they initially sunk to the bottom and then they were covered in bubbles because their surface attracts the bubbles to adhere to them. Some of them do float, but some of them stay at the bottom; while some jump up and down. I think because raisins are heavier than the 7UP, and because you are dropping them from a higher point than the 7UP initially brings them to the bottom of the glass. And then the bubbles adhere to the surface of the raisins then they can lift them up for a little bit. For those that are smaller and lighter the bubbles can keep them at the top; while the ones that are heavier are too big to stay at the top, so they go back to the bottom. The bubbles add air to make the raisins lighter in the 7UP. At the surface the bubbles pop and there isn't that force of the bubbles keeping them up anymore.

Pieces of knowledge (PK). As with observations and inferences, pieces of knowledge were assessed based on their relevance and accuracy. As evident in Table 4.5, scientists again demonstrated the highest use of pieces of knowledge: out of a total of 85 pieces of knowledge generated by all participants, 41 PK (48.2%) were produced by scientists, 24 PK (28.2%) by teachers, and 20 (23.5%) by students.

Similar to results from the CIJ demonstration, scientists tended to use significantly more PK to support their inferences and observations than students and teachers. In fact, while constructing DR explanations, students and teachers heavily used inferences to support their observations. An examination of the percentage frequencies of observations, inferences and pieces of knowledge across the three groups reveals, similar to findings from the CIJ scenario, that students and teachers produced significantly more inferences and observations than scientists.

Figure 4.25 shows that 49% of all observations were produced by students, 29% by teachers, and 22% by scientists. Additionally, 40.3% of all inferences were produced by teachers, 32.3% by students, and 27.3% by scientists. However, scientists produced significantly more pieces of knowledge than teachers and students: 48.2% of all PK were produced by scientists, 28.2% by teachers, and 23.5% by students.

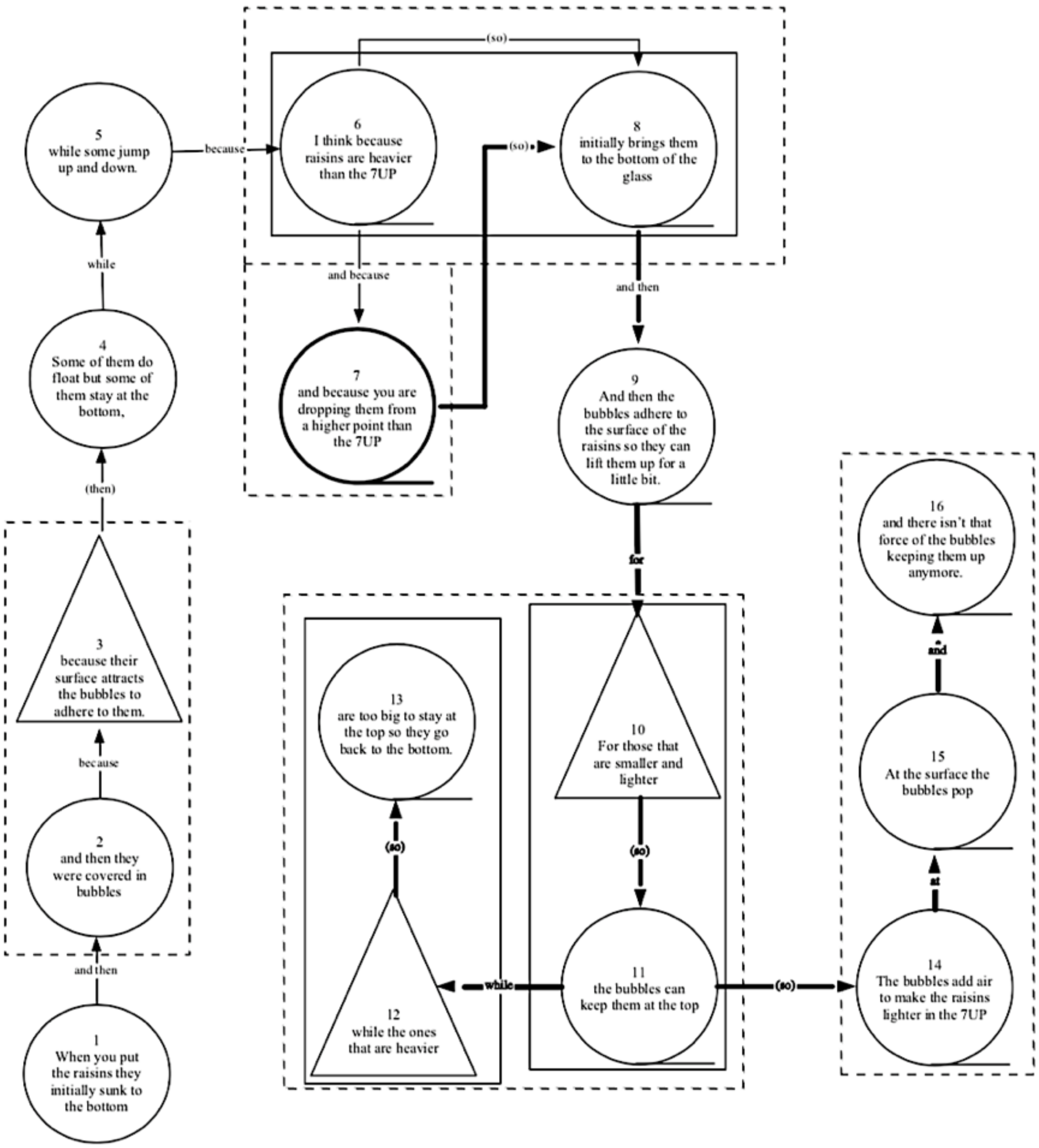


Figure 4.24 Sylvia's explanation map. An example of a scientist's map of the DR demonstration.

Additionally, three types of PK were observed among at least two of the three participating groups: accurate relevant PK, inaccurate relevant PK, and irrelevant accurate PK. What is more, scientists produced only accurate and relevant pieces of knowledge (i.e., all 41 scientist-generated PK were accurate and relevant). However, of the 24 PK produced by teachers, 79% of them were accurate and relevant, and the rest were either inaccurate or irrelevant.

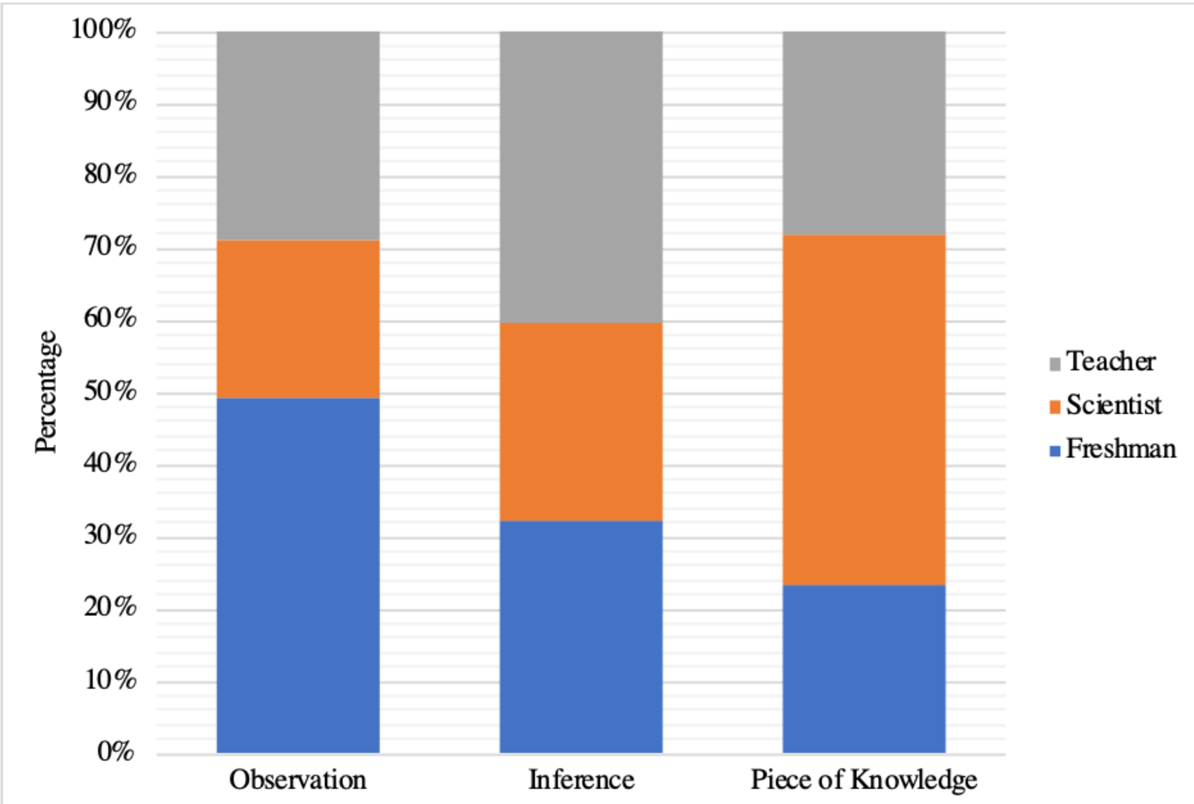


Figure 4.25 Participants' percentage frequencies of observations, inferences and pieces of knowledge in the DR demonstration across the three groups.

Similarly, of the 20 PK produced by students, 90% were relevant and accurate, and the rest were either inaccurate or irrelevant. The most common inaccurate piece of knowledge generated by students and teachers was related to the nature of the bubbles in the soda. More specifically, a considerable number of students said that the bubbles in the soda were air bubbles or made of oxygen instead of carbon dioxide. It is worth noting that this type of inaccurate piece of knowledge did not render an explanation inadequate. It only pertained to a basic inaccurate prior piece of knowledge included within an explanation.

The use of predictions in DR explanations. Unlike the CIJ demonstration, a common pattern across all participants in relation to predictions was observed when analyzing participants' DR explanations. According to the NOSE framework, predictions are regarded as statements that posit the consequences of a phenomenon prior to its occurrence. Recall that, in explanation maps, a prediction statement is included in a circle circumcised in a square (see Table 3.1 for more details). An interesting finding revealed that scientists tended to make predictions related to what would happen to the raisins after a certain period of time – more so than teachers and students. In particular, 8 of the 10 scientists, 4 of the 10 teachers, and 3 of the 10 students predicted that the raisins would all eventually sink to the bottom. What is more, all eight scientists further added reasons why they thought the raisins would eventually sink. For instance, Saul stated that "the cycle repeats. And then over time there is less gas, so eventually they will all sink," and Sara predicted that the raisins "continue doing that until which point the CO_2 is all gone out of the solution and the drink goes flat." On the other hand, students who made predictions about the DR phenomenon did not always justify their predictions. For example. Florence said that "eventually over time the raisins will come to a standstill; most of them will sink at the bottom and maybe very few will stay at the top" without clarifying why they would behave that way. Finally, teachers provided a different kind of prediction: in addition to making predictions about the behavior of the raisins after a certain period of time, 4 of the 10 teachers also made predictions related to how the raisins would behave if they were added to water instead of a soda drink. This kind of prediction, prevalent among teachers, was inferential in nature as it was followed by inferences and pieces of knowledge related to the role that the gases or the bubbles in the soda played in the DR phenomenon.

Necessary conditions (NC). Another common pattern observed across all participants while constructing DR explanations was related to necessary conditions produced by participants from all three groups. According to the NOSE framework, a necessary condition is a condition that must be present for an event to occur. Recall that, in explanation maps, statements of necessary conditions are included in a triangle (see Table 3.1 for more details). A considerable number of participants from all three groups generated necessary conditions. As evident in Table 4.5, scientists demonstrated the highest use of necessary conditions: out of a total of 37 necessary conditions, 17 NC (46%) were produced by scientists, 11 NC (30%) by students, and 9 NC (24%) by teachers. An examination of the nature of these conditions revealed interesting findings across the three participating groups. Table 4.6 presents the nature and frequency of the various necessary conditions that were generated by all participants. Results showed that there were three major necessary conditions that at least 50% of participants from at least one group generated. More specifically, these were conditions related to the surface or coating of the raisins (produced by 5 scientists, 2 students, and 1 teacher), conditions related to the bubbles at the top of the glass (produced by 6 students, 4 scientists, and 4 teachers), and conditions related to the number of bubbles on the raisins (produced by 5 scientists, 3 teachers, and 2 students). Other themes, although less commonly produced, were related to the size, shape and position of the raisins. It is worth mentioning that the type or quality of a given explanation was not affected by the existence, or lack thereof, of these conditions. In other words, providing further explanation to certain aspects of the DR phenomenon (such as why the bubbles stick to the raisins, or why the bubbles pop at the surface, etc.) did not determine the adequacy of a given explanation. Nonetheless, when such further explanation was present, structural elements were analyzed in accordance with the NOSE framework.

Hence, while generating their DR explanations, 5 of the 10 scientists included condition statements about the surface or coating of the raisins that allowed the bubbles to adhere to the raisins. For instance, Stefan stated that the "rugosity of the surface of the raisins allows for the entrapment of the bubbles or gas," Sylvia generally mentioned that "their [the raisins'] surface attracts the bubbles to adhere to them," and Stanley simply described that the surface of the raisins is wrinkly and "bubbles attach easier on wrinklier surface." Finn and Franco were the only two students who further explained why the bubbles adhere to the raisins. Both of them mentioned statements about the wrinkly surface or coating of the raisins. Finally, of the 10 participating teachers, Trevor went in-depth about explaining what he thought was the reason for the bubbles to stick on the raisins. Trevor explained:

> And the reason why the bubbles stick to the raisins is some sort of attractive force between the bubbles and the surface of the raisin. There is some intermolecular force between them; it is some sort of dispersion force because CO_2 is nonpolar and it doesn't have dipole-dipole. So there is some force that makes them stick to the raisins.

In addition to necessary conditions related to the surface of the raisins, some participants further provided condition statements about the location where the bubbles pop: at the top or at the surface of the cup. In particular, 6 of 10 students, 4 of 10 scientists, and 4 of 10 teachers explicitly specified the location at which the bubbles popped. What is more, participants from only the scientists and the teachers group further explained why. While students simply pointed out that the popping of the bubbles occurred at the surface, teachers and scientists further explained surface tension as the main reason for the bubbles popping at the top. For instance, Thomas explained that "when the raisins with the bubbles reach the top they reach the surface and break through the surface tension."

Table 4.6

The Nature and Frequency of Participant-Generated Necessary Conditions in the Dancing Raisins Phenomenon Among All Participants

	Students	Scientists	Teachers
Necessary Condition			
Surface of the raisins	2	5	1
At the top	6	4	4
Number of bubbles on the raisins	2	5	3
Size or shape of the raisins	1	2	1
Position of the raisins	--	1	--

Finally, a major necessary condition that was observed across all groups – though with varying degrees of occurrences – was related to the number of bubbles on the raisins. While participants explained the floating of the raisins due to the bubbles adhering to the raisins, scientists were more specific in explaining the relation between the shifting of the raisins and the number of bubbles on the raisins. For instance, Samantha pointed out that "sometimes all the leftover bubbles are not enough to support it [the raisin], and then the raisins sink again"; while Sophia stated that "the raisins that were more exposed to the bubbles were then allowed to float back up because they had enough bubbles latch onto them or around them."

The nature of explanatory connections or big ideas made. Similar to findings from the CIJ demonstration, when constructing DR explanation maps, it was evident that several structural elements together made explanatory connection or a big idea. An examination of the structural elements produced by all participant groups to make explanatory connections in the DR scenario revealed that scientists and teachers made more explanatory connections that students (a finding different than that of explanatory connections of the CIJ demonstration). In fact, teachers and scientists made more connections (51) than students (41 connections). Additionally, scientists' and teachers' explanatory connections included more structural elements than those made by students. What is more, even though all participants' connections included more inferences than pieces of knowledge, students' and teachers' connections included significantly more inferences than pieces of knowledge than did scientists' connections.

As evident in Figure 4.26, 32% of the total structural elements produced by scientists to make explanatory connections were inferences, followed by pieces of knowledge (27%), laws and lawlike statements (19%), necessary conditions (11%), and predictions (6%), with only 5% being observations. On the other hand, teachers' explanatory connections were made of various structural elements in which 58% of these elements were inferences, followed by pieces of knowledge (13%), laws and lawlike statements (12%), observations (10%), and necessary conditions (5%0, with only 2% being predictions. Finally, students' explanatory connections included various structural elements where 54% of these elements were inferences, followed by observations (20%), pieces of knowledge (12%), laws and lawlike statements (8%), and necessary conditions (5%), with only 1% being predictions.

In summary, an analysis of explanation maps in accordance with the NOSE framework indicated that participants tended to use certain structural elements when building their explanations of the DR phenomenon. Even though all participants heavily used inferences while constructing their DR explanations, scientists' ratio of the use of inferences to pieces of knowledge was significantly less than that of the ratios of the two other groups. In addition to the fact that scientists generally produced less inferences and observations, they demonstrated the highest ratio of accuracy and relevancy of these elements as compared to those produced by teachers and students. Finally, it was observed that the use of more necessary conditions to provide further explanations to various aspects of the DR phenomenon was not necessarily an indication of a more adequate scientific explanation.

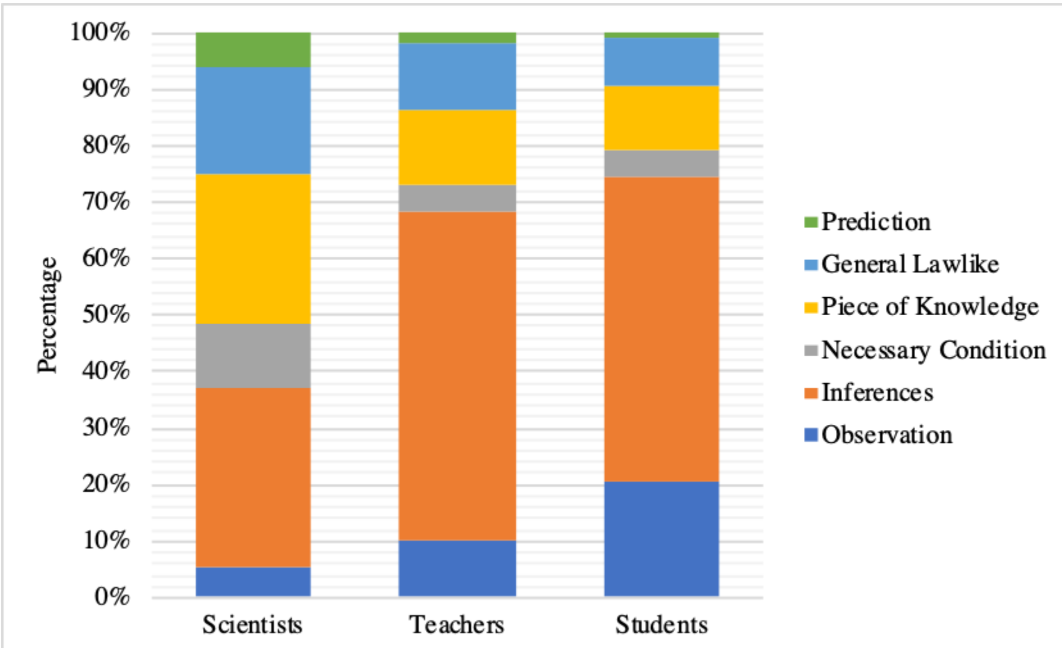


Figure 4.26 Participants' explanatory connections in the Dancing Raisins demonstration. The structural elements are based on percentage frequencies.

Basic features of explanation maps within groups. Several major aspects emerged when examining explanation maps. These were: (a) presence of the aforementioned structural elements, mainly, the use of inferences, pieces of knowledge and necessary conditions; (b) use of causal links and causal connections; and (c) types and quality of the explanations constructed. The following sections present these results within each participant group.

Practicing scientists' explanations. Table 4.7 shows that a total of 174 statements were observed in scientists' DR maps; of those 65 (37.3%) were inferences, 41 (23.5%) were pieces of knowledge, 29 (16.7%) were laws and lawlike statements, 17 (10%) were necessary conditions, 13 (7.5%) were observations, and 9 (5%) were predictions. Of the inferences, scientist explanation maps showed mostly the use of accurate relevant inferences (75%) and the use of only accurate and relevant PK (100%).

First, it is worth mentioning that when asked all participating scientists said that they had not seen or done this particular experiment before. However, two of the 10 participating scientists said that they had seen something similar to it, and two others compared the experiment with phenomena from everyday life. In particular, Sara and Saul mentioned that this experiment was similar to the Diet Coke and Mentos Eruption demonstration that they had seen a video of. Sophia mentioned that this experiment (mainly the bubbles forming on the raisins) resembled the phenomenon of floating in a swimming pool; while Stella said that the experiment reminded her of a similar phenomenon when she put fruits in champagne.

Table 4.7

Frequency of the Main Structural Elements of the Dancing Raisins Phenomenon Within Groups

	Scientists		**Teachers**		**Students**	
Structural Elements	*f*	%	*f*	%	*f*	%
Observation	13	7.5	17	10	29	18
Inference	65	37.3	96	55	77	47.5
Piece of Knowledge	41	23.5	24	14	20	12.3
Necessary Condition	17	10	9	5	11	7
Law and lawlike statement	29	16.7	22	12.6	18	11
Prediction	9	5	6	3.4	7	4.2
Total	**174**		**174**		**162**	

An overall examination of scientists' maps indicated that all scientist-generated explanations were, in fact, scientific (see Table 4.8). Of these 10 scientific explanations, seven were adequate, two were mostly adequate, and one was only partially adequate. What is more, all of the 10 explanations included causal links of some sort: only 1 of the 10 included simple cause-effect relationships thus producing a causal explanation, while the remaining nine included causal connections in the form of casual processes and causal interactions that lead up to the phenomenon to be explained. Recall Sylvia's DR explanation (see Figure 4.24): Sylvia used simple cause-effect relationships to explain the DR phenomenon, thus producing a partially adequate causal explanation. Of the remaining nine CM explanations, eight were subsumed under natural laws and lawlike statements, thus producing CMDN explanations, and one was subsumed under a probabilistic/statistical type law, producing a CMIS explanation. However not all of these nine explanations were adequate: 6 of the 8 CMDN explanations were adequate, the one CMIS was also adequate, but two were mostly adequate. Sophia and Stella were two scientists who provided mostly adequate CMDN explanations. Even though Sophia used density-related inferences to explain the behavior of the raisins via causal-mechanistic processes, she interchangeably used laws and lawlike statements that related heaviness (or weight) to the sinking or floating of the raisins. Sophia's full explanations with its corresponding explanations amps is presented in Figure 4.27:

> Because you are adding mass to the cup, you are going to displace some of the mass in the cup; and this where we saw the initial pushing up of the bubbles in the liquid. The raisins sink to the bottom because they are more dense than the liquid. The raisins are solid, and they are more compact, and they are heavier than the soda so they sink. Density is a property that every molecule or a compound has. It describes their mass to

volume quantity. And the raisins that were more exposed to the bubbles, were then allowed to float back up because they had enough bubbles latch onto them or around them. That gave them buoyancy. Kind of like the same phenomenon when you are floating in a swimming pool. If you position yourself in a certain way you are more buoyant than if you just dive down. Also some of them are positioned in such a way they could float with the extra air bubbles around them. But the other ones are compacted and maybe stuck together or to the walls of the cup. So they have less surface exposed to the bubbles to latch on them, so they stay at the bottom. The bubbles are actually made of CO_2. And gas tends to be lighter than the liquid, so it tends to float up. And that's why you see the bubbles float up. And so if they are already floating up and can attach to the raisins then they can help them come up.

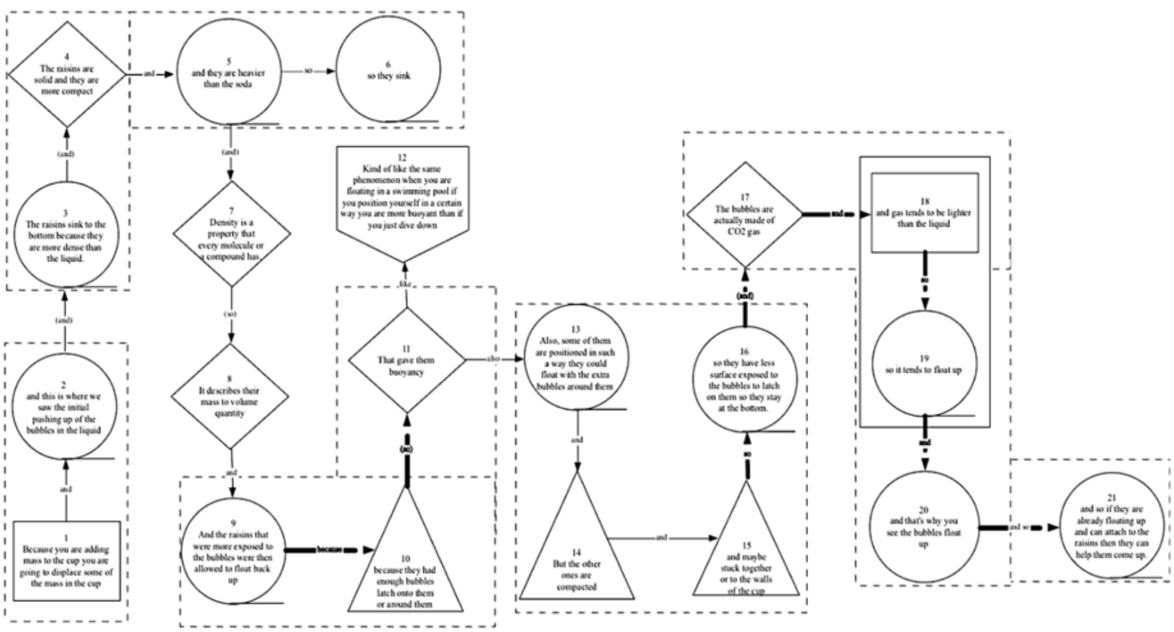


Figure 4.27 Sophia's explanation map. An example of a mostly adequate CMDN scientist DR explanation map.

Table 4.8

Nature and Quality of the Dancing Raisins Explanations by Explanation Type Generated by Practicing Scientists (S), High School Science Teachers (S), and Freshman Students (F)

	DN			Causal			CDN			CMDN			CMIS			Total (By Quality)
	S	**T**	**F**	**S**	**T**	**F**	**S**	**T**	**F**	**S**	**T**	**F**	**S**	**T**	**F**	
Adequate	--	--	--	--	--	--	--	--	--	6	5	4	1	--	--	**16**
Mostly Adequate	--	--	--	--	--	--	--	--	--	2	3	1	--	--	--	**6**
Partially Adequate	--	--	--	1	--	1	--	--	1	--	--	--	--	--	--	**3**
Inadequate	--	--	1	--	--	2	--	--	--	--	2	--	--	--	--	**5**
Total (By Type)	**1**			**4**			**1**			**23**			**1**			**30**

High school science teachers' explanations. Ten high school science teachers generated DR explanations. Table 4.7 shows that a total of 174 statements were observed in teachers maps; or those 96 (55%) were inferences, 24 (14%) were pieces of knowledge, 22 (12.6%) were laws and lawlike statements, 17 (10%) were observations, 9 (5%) were necessary conditions, and only 6 (3.4%) were predictions. Of the inferences, teacher explanation maps showed mostly the use of accurate and relevant (83.3%) inferences, and all lawlike statements in teacher explanation maps were general deterministic laws and natural regularities.

Unlike scientists, teachers' familiarity and prior knowledge related to this experiment was evident. Five teachers said that they had seen and/or done this experiment before, although some said it was done differently. For example, Tammy said that she had done this experiment but instead had asked the students to guess what the drink was; while Tucker said that he had done this experiment but collected students' predictions of what they thought would happen to the raisins if left soaking in for a few days. In fact, Tucker said that he had a beaker with raisins and soda in his science room that had been sitting there for a day. Both Tucker and Tammy taught high school biology. Tarra, who was a novice high school teacher with one year of teaching experience, said that she had seen this experiment done in one of her science methods classes during her Master in Science Education. Additionally, Thomas and Trevor said that they had seen this experiment somewhere but did not remember where. Thomas further compared the experiment with someone "wearing a floatation device in water". Similarly, Todd compared the bubbles sticking on the raisins with "floaties" in a swimming pool. Tanya and Tyson reported that they had not seen this experiment before, but that they had seen something similar, particularly Diet Coke and Mentos Eruption demonstration. Finally, Tracy and Tina said that they had never seen or done something like this experiment before.

Hence, an overall examination of teachers' explanation maps indicated that all of teacher-generated explanations were in fact scientific. As evident in Table 4.8, teachers' DR explanations revealed less variation in terms of type than their CIJ explanations. In particular, all 10 teacher-generated explanations were CMDN: in all of their DR explanations teachers tended to include causal connections in the form of causal processes and interactions that lead to the phenomenon-to-be-explained subsumed under natural laws and lawlike statements. However, not all these CMDN explanations were adequate: 5 of the 10 CMDN explanations were adequate, three were mostly adequate, and two were inadequate CMDN explanations. Todd provided an adequate CMDN explanation in which he explained the DR phenomenon using only relevant and accurate structural elements. Todd's full explanation and its corresponding explanation map are presented (see Figure 4.28):

> The raisins are going up and down, and now it is a cycle of the popping of the bubbles and the forming of new ones. The density of the raisins is greater than the density of the fluid, so initially they sink to the bottom. This is due to their buoyant force. You can compare the displaced volume of weight of the water vs the weight of the raisin to figure out which is denser. If it is denser, then it is going to sink, unless it has bubbles. So as the bubbles sink, they grow. And we have more coming out of the solution. And as they get to the top, the bubbles that reach the surface pop and the raisins sink back to the bottom. Then they grow more bubbles, and the cycle repeats until there are no more bubbles. The bubbles are creating a force upwards that is attached to the raisins. It is like if you were under water and there was a floatie at the bottom and you grab a hold of it, it will make you float to the top. When they reach to the top they pop due to surface tension. After a few hours the raisins will probably all stay at the bottom, since the raisins are more dense than the solution. If we draw a force diagram on the raisin, there is gravity pulling it down, and that force is the biggest force, so it goes to the bottom.

Recall that both Tucker and Tina (Figures 4.22 and 4.23) provided inadequate DR explanations. They both included mostly irrelevant and inaccurate structural elements to explain the behavior of the raisins. Both Tucker and Tina explained that the density of the raisins themselves changed. Tucker explained that the raisins *oscillated* in the soda due to membrane exchange and diffusion that in turn lead to density changes; whereas Tina explained that the raisins used a *convection cycle* as they got filled with gas.

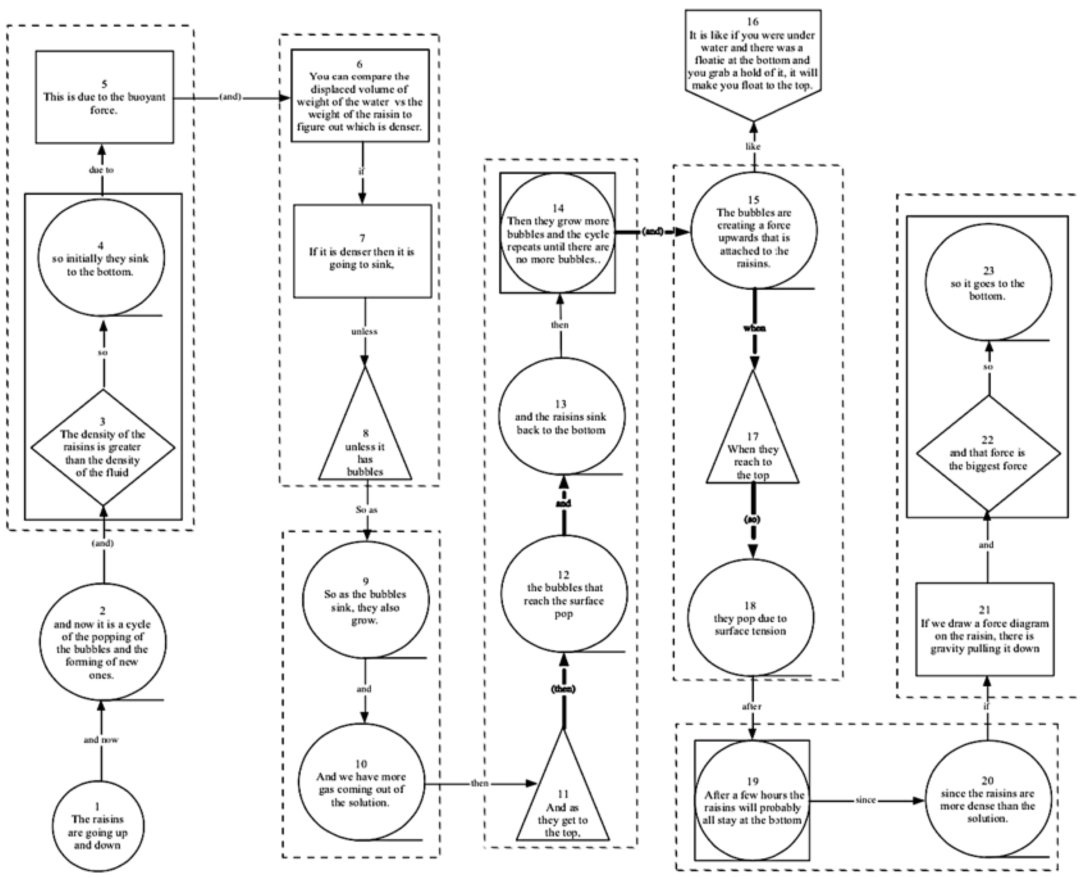


Figure 4.28 Todd's explanation map. An example of an adequate CMDN teacher DR explanation map.

On the other hand, Tammy, Tarra and Tyson provided mostly adequate CMDN explanations that included mostly accurate and relevant structural elements. For example, Tarra adequately explained the behavior of the raisins relating density to floating and sinking. However, in her explanation she also included inaccurate statements such as an inference she made that the raisins *chemically reacted* with the bubbles, in addition to accurate statements that the raisins had bubbles on them. Finally, both Tammy and Tyson used mostly accurate and relevant structural elements to explain the behavior of the raisins; however, they also included statements relating the weight (heaviness and lightness) to the raisins sinking and floating.

College freshman students' explanations. Ten college freshman students produced explanations of the DR. Table 4.7 shows that a total of 162 statements were observed in student maps; of those 77 (47.5%) were inferences, 29 (18%) were observations, 20 (12.3%) were PK, 18 (11%) were laws and lawlike statements, 11 (7%) were necessary conditions, and only 7 (4.2%) were predictions. Of the inferences and PK, student explanation maps showed that they were mostly relevant and accurate (73% and 80% respectively). However, the number of PK students produced was roughly one quarter that of their inferences.

When asked if they had seen or done the DR demonstration or seen something familiar to it, all participant students said that they had not seen this particular experiment before, but five said that this experiment was similar to the Diet Coke and Mentos Eruption demonstration that they had seen videos of. Additionally, two freshman students compared the bubble formation on the raisins (and the floating of the raisins) with humans floating in a pool. Filip stated that "it is like when we wear life jackets and jump into the ocean. The life jackets keep you afloat because they have a lot of air in there. So the air bubbles here are trying to act as life jackets for these raisins"; whereas Fidel stated: "Just like if you were to have a floatie in a pool, something filled with air helps you float to the top. Similarly, the bubbles that are on the raisins act like floaties that help raise it up to the top".

An overall examination of students' explanations indicated that all student-generated explanations were in fact scientific. Students produced more variations of explanations by type and quality compared to the other two groups. As evident in Table 4.8, 4 of 10 explanations were adequate, one was mostly adequate, two were partially adequate, and three were inadequate. What is more, 5 of 10 explanations included causal connections in the form of causal processes and interactions that lead up to the phenomenon-to-be-explained that were subsumed under natural laws, thus producing CMDN explanations. Four of these CMDN explanations were adequate, while one was mostly adequate. Faith produced a mostly adequate CMDN explanation that included mostly accurate and relevant structural elements. However, in her explanation Faith also included some inaccurate elements. More specifically, Faith explained that the raisins sink because they were “not buoyant” instead of comparing the buoyant force with the weight of the raisin.

Simple cause-effect relationships were identified in four of the students’ explanations. In particular, Felicia and Florence produced inadequate causal explanations. Felicia inadequately explained the sinking of the raisins because of their mass, and explained the floating of the raisins due to a chemical reaction between the soda gases and the raisins. Felicia explained (see Figure 4.29):

> Once you put the raisins in they all went to the bottom. They go down right away because of the mass. Then they started jumping around. The gases from the 7UP made the raisins jump around up and down. So, it is the gases in the soda that is making them react somehow. It is some kind of a reaction into the actual raisin. It is not water. It is plain liquid. Since soda has gases. So, it is making them do something different. Some of them sink all the way at the bottom and stay there, and then the rest came up and went down.

Franco produced an inadequate DN explanation in which he explained the DR phenomenon as a consequence of irrelevant and inaccurate laws and lawlike statements. First, Franco related sinking and floating to the mass of the raisins. Franco did cite density once, but in

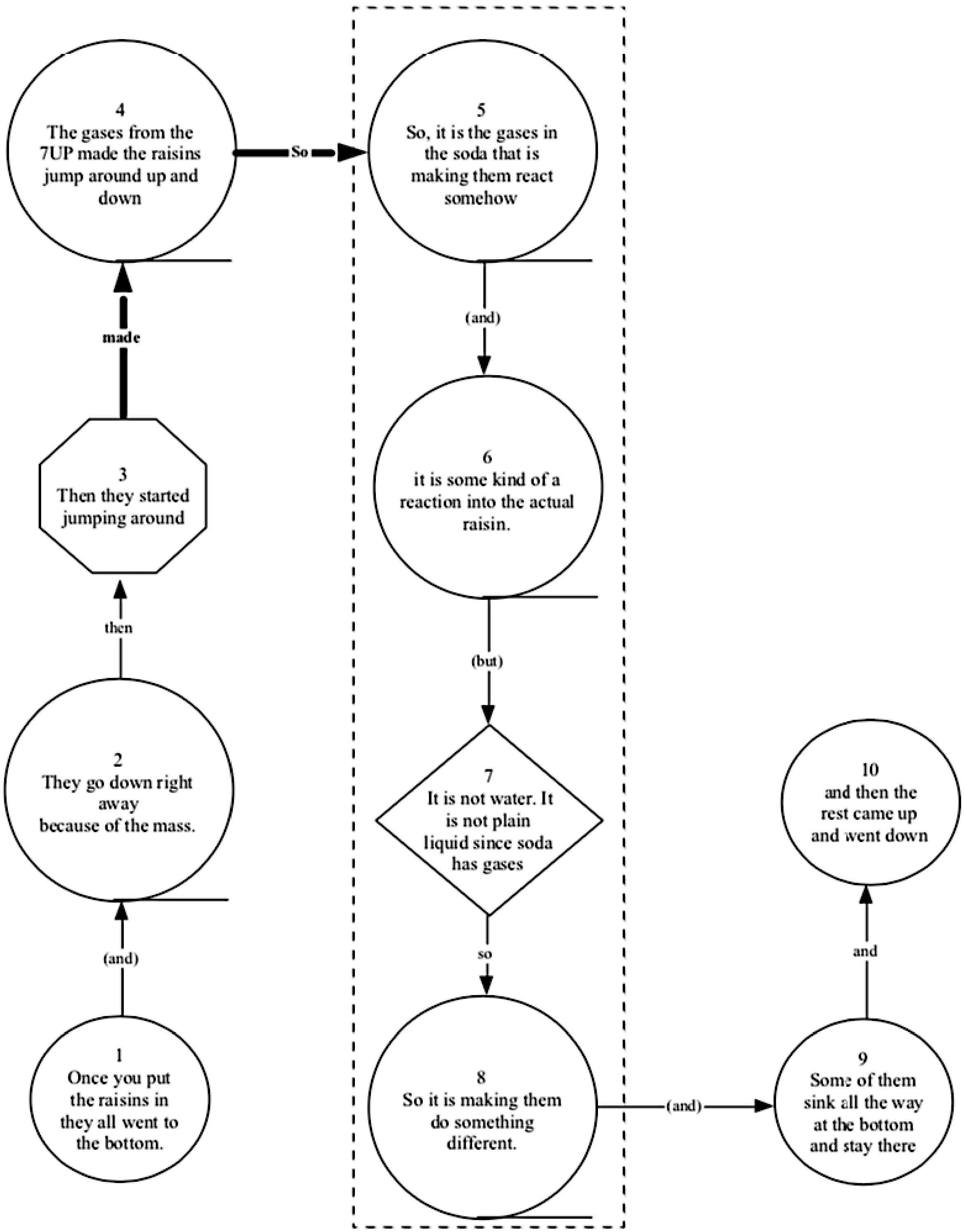


Figure 4.29 Felicia's explanation map. An example of an inadequate causal DR explanation map produced by a freshman student.

a vague and general lawlike statement: “They first sink because of the whole density of the raisin thing,” but then when he elaborated, he only used terms such as heavy and light to explain the sinking or the floating. Furthermore, Franco considered the bubbles to be oxygen, and cited irrelevant PK and inferences about the density of the oxygen and the density of water, thus producing an inadequate set of laws and lawlike statements.

The two remaining student-generated explanations were partially adequate; one was a partially adequate causal explanation and the other was a partially adequate CDN explanation. Farrah explained the sinking of the raisins “because of their weight”, and the floating of others “because they were lighter”. She also explained that “they floated because of the bubbles; the air in the bubbles raised them up” thus producing only partially adequate causal explanation. Finally, Filip explained the behavior of the raisins as a consequence of Archimedes principle. But when explaining the sinking of the raisins, he too, stated that they sank to the bottom because of their weight. He then explained the causal link between the bubbles and the floating of the raisins. Finally, he used a mix of accurate and inaccurate structural elements that rendered his explanation only partially adequate CDN. Filip’s full explanation map is presented in Figure 4.21.

Comparison of the Three Groups. Of the 30 participant-generated scientific explanations of the dancing raisins phenomenon, 16 were adequate; seven of which were produced by scientists, five by teachers and four by students. Similar to the case of the CIJ, teachers were more familiar with the DR than the two other groups, and some had previously seen and done the demonstration. While generating DR explanations, causality was more common in participants’ explanations across the three groups than it was among participants’ CIJ explanations.

Hence, an examination of all participants' types of DR explanations shows that CMDN explanation type was the most generated among the three group (23 out of the 30 explanations were CMDN), followed by four Causal explanations, one DN explanation, one CDN explanation, and one CMIS explanation. Hence, the use of causal links and causal connections (simple and multiple connections) was evident as 29 of the 30 participant-generated DR explanations were causal. As would be expected, and similar to the CIJ scenario, scientists demonstrated the highest use of multiple causal connections subsumed under laws and lawlike statements; while teachers and students used more simple cause-effect relationships to indicate causality.

Similar to the CIJ scenario, overall, inferences and pieces of knowledge were the major structural elements observed in participant maps with differences in the frequency and ratio of each per group, their relevance and accuracy. Despite the fact that scientists used more inferences than pieces of knowledge while constructing their DR explanations (a different finding than what was observed in the CIJ), they still used significantly more PK than students and teachers. What is more, the ratio of inference to PK was strikingly different between scientists on the one hand, and teachers and students on the other hand: students (77 inferences and 20 PK) and teachers (96 inferences and 24 PK) produced roughly four times more inferences than PK; whereas scientists produced one and a half more inferences than PK (65 inferences and 41 PK).

Finally, common patterns were observed across all participants in relation to various structural elements of both scenarios: the Dancing Raisins scenario, and the Burning Candle scenario. The following section presents these basic and common features of explanation.

Basic features of explanation maps between the two scenarios.

An examination of all 60 final explanations maps (30 CIJ explanations and 30 DR explanations) revealed common patterns across all participants in relation to the use observations, inferences and pieces of knowledge, the use of causal links and causal connections, the nature and variety of laws and lawlike statements, and the types of explanations produced.

Observations, inferences and pieces of knowledge. In both scenarios, students and teachers demonstrated higher use of observations and inferences compared to scientists. What is more, while constructing CIJ and DR explanations scientists produced significantly less inaccurate and/or irrelevant observations and inferences than students and teachers. Additionally, in both scenarios, similar numbers of observations were produced in each group: scientists produced 12 CIJ observations (and 13 DR observations), teachers produced 22 CIJ observations (and 17 DR observations), and students produced 28 CIJ observations (and 29 DR observations). On the other hand, significantly more DR than CIJ inferences were produced across all three groups: scientists produced 29 CIJ inferences (and 65 DR inferences), teachers produced 54 CIJ inferences (and 96 DR inferences), and students produced 57 CIJ inferences (and 77 DR inferences). Finally, while scientists and teachers produced similar numbers of pieces of knowledge in both CIJ and DR, students produced almost twice as much DR PK than CIJ PK: scientists produced 38 CIJ PK (and 41 DR PK), teachers produced 25 CIJ PK (and 24 DR PK), while students produced 11 CIJ PK (and 20 DR PK). It is clear that the CIJ phenomenon is more complex than the DR. Interestingly enough, however, participants did not produce more pieces of knowledge while constructing their CIJ explanations. In fact, as mentioned earlier, students produced significantly more DR PK. This is was not the case with the numbers of laws and lawlike statements, however – as will be elaborated later in the chapter.

Given the broader scope of knowledge, it was expected that teachers and scientists would draw more on prior knowledge than students, and that was evident when the former two groups would link their explanations in both scenarios to supporting relevant inferences, pieces of knowledge, and lawlike statements instead of sensory observations. What is more, teachers and students were probably less confident about their knowledge base compared to scientists. Thus, students and teachers could have been more hesitant in adding more pieces of knowledge to support their explanations. Hence, they tended to use more deterministic statements such as "it is a law of nature"; "the nature of entropy tells us that"; and "that follows the laws of thermodynamics" as these were examples of statements produced by teachers and students.

The use of causal links and causal connections. Another significant finding across the two scenarios was related to the use of causality in participant-generated final explanations. In particular, the use of simple and multiple causal links was significantly higher in DR explanations than in CIJ explanations. In both scenarios, scientists exhibited the highest use of multiple causal links than teachers and students. Additionally, while students and teachers tended to use more simple cause-effect relationships to explain the CIJ phenomenon than scientists, only students used mostly simple cause-effect relationships to explain the DR phenomenon. In other words, while causal mechanisms were not very common in teachers' CIJ explanations, they were more common in their DR explanations.

The nature and variety of laws and lawlike statements. In both scenarios, laws and lawlike statements used were mainly deterministic in nature. It can be argued that the nature of both CIJ and DR scenarios allowed for such type of laws and lawlike statements. Participants used mostly natural regularities and general laws to explain both the CIJ and DR phenomena. Of all 30 participants, one participating scientist, Sam, tended to use more probabilistic and statistical versions of general laws. In addition, based on the nature of the phenomenon-to-be-explained, it was evident that there was a significantly more variety of the laws and lawlike statements employed in all participants' explanations in CIJ explanations than in DR explanations. More specifically, an examination of participants' CIJ explanations revealed at least six different scientific laws (e.g., pressure difference, pressure equilibrium, Ideal Gas Law, expansion of a gas, contraction of a gas, etc.). However, this was not the case of the DR phenomenon: scientific laws and lawlike statements employed in participants' DR explanations were mainly related to density, buoyancy, weight/mass/gravity, and Archimedes Law. This finding highlights the different complexities of the two phenomena. However, while this complexity difference was not reflected in the use of pieces of knowledge, it was evident in the variety of the use of laws and lawlike statements. When explaining the CIJ, several scientific laws need to be included for an adequate explanation. This is mainly due to the nature of the phenomenon itself since it cannot be explained through oxygen consumption alone. On the other hand, a density-related explanation is adequate to explain the behavior of the raisins in the DR demonstration. Thus, this highlights the NOSE framework emphasis on the pragmatic nature of scientific explanation: what counts as an adequate explanation is not necessarily determined by the number of structural elements, but rather by the nature (accuracy and relevance) of these elements.

Unlike other frameworks that have been used in science education research studies to analyze students' explanations, NOSE framework enables researchers to make explicit the differences between explanations and non-explanations in science in addition to the various structural elements that make up a given explanation. Furthermore, NOSE framework examines the accuracy and relevance of structural elements in a scientific explanation without disregarding the content is which it is produced. In particular, the relevance of inferences and pieces of knowledge, the logic of laws and lawlike statements, and the accuracy of causal links all depend on the context in which the explanation is produced. Following the NOSE framework, a piece of knowledge in an explanation might be analyzed relevant to a phenomenon A but irrelevant to another phenomenon B.

Perceptions of Scientific Explanation

What is Scientific Explanation? This study aimed to examine how participants perceived explanations within a scientific context. Results are presented in sections according to the different participant groups.

Students. The majority of student participants (8 of 10 students) considered scientific explanation be an answer to a scientific *why*-question. Student participants identified scientific explanation as one that provides a reason as to why something happens in the world. In addition, half of the students considered scientific explanation to be based on observations. More specifically, students believed that scientific explanation must include all observations related to the phenomenon to be explained. Moreover, four students considered scientific explanation to be objective or "always true." This meant that a scientific explanation must not leave "much up to someone's assumptions" (i.e., not based on a person's opinion and has no room for misinterpretations), uses "objective terminology" (i.e., uses scientific terms), and one that gives you "no option but to agree with it." In addition, four student participants viewed scientific explanation as a statement, or a group of statements supported by scientific theories, laws and principles. Finally, four student participants emphasized that scientific explanation is based on experimental evidence and scientific facts. In other words, students believed that scientific explanation should present an answer to a *why*-question based on scientific facts. What students considered facts included existing scientific knowledge, knowledge resulting from scientific experiments (e.g., laws, theories), and data. For example, Florence considered a scientific explanation to be "an explanation to a problem that you have tested and gotten results from it and come to a conclusion as to how and why it works."

Minor themes or features of scientific explanation were emphasized by a few students. For instance, two students noted that a scientific explanation needed to be supported by mathematical equations or mathematical proof. Only one student emphasized that scientific explanation must follow the scientific method; while another student considered scientific explanation to be a series of cause-effect relationships based on observable and measurable entities.

In summary, the majority of students considered scientific explanation to be a statement or a group of statements that answers why questions while also describing what happened (for some in an objective unbiased manner). Table 4.9 sums up the generated features from student transcripts with accompanying frequencies regarding what student participants considered scientific explanation to be.

Teachers. Six of 10 teacher participants emphasized that scientific explanation should provide understanding and help make sense of the natural world. Similar to students, three teachers stated that scientific explanation provides an answer to a scientific *why*-question (i.e., gives a reason why something happened rather than what happened). In addition, five teachers also viewed scientific explanation as a statement or a group of statements that contains scientific principles, scientific laws, and/or scientific facts that are supported by observations and/or mathematical equations.

Teachers also noted that evidence in scientific explanation should be empirical or quantitative and must clearly support observations and conclusions. However, four teachers emphasized that scientific explanation must take the form of a claim-evidence-reasoning.

Table 4.9

Students' Features of a Scientific Explanation

Feature	Illustrative excerpt	*f*
A scientific explanation is an answer to a *why*-question	Faith: A scientific explanation is a recap of what was going on along with details of why it is going on.	8
A scientific explanation is based on observations (i.e., it includes all observations related to the phenomenon-to-be-explained)	Felicia: A scientific explanation is a description of the observations that you have encountered and the best way that you believe you can explain what you just saw.	5
A scientific explanation is objective or always true (i.e., it does not leave much up to someone's assumptions, there is no room for misinterpretations, and uses objective terminology)	Franco: A scientific explanation is one that if you are explaining what happened it explains entirely what happened and gives you the whole breakdown. There is no room for error or misinterpretation.	4
A scientific explanation is supported by scientific theories, scientific laws and scientific principles.	Fidel: A scientific explanation is one that explains a phenomenon in detail given some sort of property or physical or theoretical principle or law.	4
A scientific explanation is based on experimental evidence and scientific facts.	Florence: A scientific explanation is a solid explanation to some kind of phenomenon that explains with the use of facts, descriptions to explain something to the point where there is no option but to agree with it.	4
A scientific explanation is supported by mathematical proof and mathematical equations	Filip: A scientific explanation is something that introduces a new concept using some existing concept or principle that the student already knows and trying to build off of it and supporting with experiments, observations and some mathematical explanations.	2
A scientific explanation must follow the scientific method	Flynn: A scientific explanation should provide reasonable basis with evidence and follow the scientific method. So you look at the experiment and try to describe what happened in our own words and why did it happen.	1
A scientific explanation is a series of cause-effect relationship	Frederick: A scientific explanation is like a series of cause and effect relationship that is based on something that is observed and measured.	1

They believed that scientific explanation contained elemental features significant to its quality. In other words, they considered scientific explanation to be made up of composite elements whose qualities should be empirically testable and measurable. These testable qualities targeted the verifiability and reproducibility of the elements that make up an explanation; mainly, hypothesis, evidence and a conclusion. What is more, three teachers considered scientific explanation to be objective or true. Like students, this meant that scientific explanation must not be based on assumptions, must not be subjective, and must be *always* true. In fact, 3 of 10 teachers believed that scientific explanation is mainly a process that follows the "scientific method." As such, a scientific explanation is considered to be "true", and it involved the explainer in observation, hypothesizing, data collection, testing, analysis, and making conclusions. Trevor for example, explained this feature with an elaboration on the elemental aspects contained within a scientific explanation:

> A scientific explanation is one that uses the scientific method: you make observations, you put the hypothesis of the reasons why you think something is occurring and then you explain it using reference to your observations and then you come up with a conclusion.

Another major feature unique to participant teachers was emphasized by six teacher. This was mainly that a scientific explanation is context-dependent; that is, an explanation depends on the learner's prior knowledge and that while it is not absolutely true, it has to be at the "appropriate level of correctness," which speaks to the pedagogical lens of science teachers. Table 4.10 sums up the major and minor features generated from teacher transcripts with accompanying frequencies of what teachers considered to be a scientific explanation.

Scientists. Almost all scientists (9 of 10) perceived scientific explanation as a statement or a group of statements that are largely based on scientific knowledge. Additionally, they

Table 4.10

Teachers' Features of a Scientific Explanation

Feature	Illustrative excerpt	*f*
A scientific explanation provides understanding (i.e., it makes sense of the natural world).	Tammy: A scientific explanation is a way to explain a natural phenomenon to make sense of the world.	6
A scientific explanation is context-dependent (i.e., it depends on the learner's prior knowledge)	Thomas: A scientific explanation has to be the appropriate level of correctness; it cannot be false but never going to be completely true. And that depends on who you are explaining to.	6
A scientific explanation contains scientific principles, scientific laws, and/or scientific facts that are supported by observations and/or mathematical equations.	Tanya: A scientific explanation is connecting what you see to the reason behind it using scientific laws and mathematical equations.	5
A scientific explanation is of the form of claim-evidence-reasoning (i.e., it contained elemental features significant to its quality)	Tarra: You have to have some sort of evidence that could be based on an observation or some sort of mathematical analysis. And you have to provide some type of reasoning along with that evidence in order to explain how you're understanding a certain phenomenon, and why that is or is not happening.	4
A scientific explanation is an answer to a *why*-question.	Tracy: A scientific explanation explains what goes on around you (…) in order to understand or tell somebody what and why something is happening. (table continues)	3
A scientific explanation is objective or always true (i.e., it must not be based on assumptions, must not be subjective, and must be always true)	Tyson: A scientific explanation clearly has correct claim and evidence. It has to have correct and clear evidence to back it up. So, it has to be accurate.	3
A scientific explanation must follow the scientific method.	Tina: My mind immediately goes to the scientific method. A scientific explanation is an explanation that explains a natural phenomenon using the scientific method. Because there are other explanations out there that might be valid to some people but don't follow the scientific method therefore they cannot be scientific explanations.	3

emphasized that scientific knowledge plays a role in connecting observations with scientific laws, and they believed that although they play an important role in an explanation, a scientific explanation *cannot* be solely reduced to them.

Most scientists (7 of 10) emphasized that scientific explanation exhibits a logical flow or is an internal consistent set, of scientific ideas (i.e., it shows logical coherence). This meant that throughout an explanation, logical connections between observations, laws and data must be explicit and clear. For example, Selena clarified what is meant by logical coherence as one important feature of the composition of a scientific explanation:

> [A scientific explanation] is taking all of the ingredients of a phenomenon and arranging them in a way that you can follow according to some internal sense of logic. I am hedging on saying being objective or reasoning in a consistent sense because I think that's a little bit up for interpretation. But it should have some internal consistency to the person who is talking about it. If you are trained as a scientist, you are taught a certain order in which you try to push the pieces and build on existing things that you already know.

In addition, 6 of 10 scientists believed that scientific explanation includes elemental aspects significant to its quality. However, unlike teachers, the qualities for scientists were more extensive and included testability, predictability, verifiability, reproducibility, and rationality. Five scientists also noted that scientific explanation must make its point by including all relevant information; and three scientists further added that a scientific explanation must dismiss irrelevant information.

Six of the participant scientists did not perceive of scientific explanation as absolutely true or objective. However, they believed scientific explanation to be *statistically* or *probabilistically* true. In other words, according to these scientists an explanation should be true a *reasonable* number of times. One scientists, Stefan, however, emphasized the objective nature of a scientific explanation that it "depends not on subjective thought or opinion but rather observable and repeatable facts." Furthermore, and similar to teacher participants, 4 of 10 scientists believed that scientific explanation provides an understanding of the natural world. Finally, four scientists also believed that scientific explanation should follow the scientific method.

Clarity and comprehensibility were also seen by two scientists as essential features of a scientific explanation. Only two scientists emphasized that a scientific explanation is context-dependent and that it has to "connect with a given audience." Finally, two scientists also viewed a scientific explanation to be dependent on the field (hard *versus* softer science, or physics, chemistry or biology). For example, Saul emphasized how explanations in physics are different than those in chemistry or biology:

> Saul: [A scientific explanation] depends on the field. In physics it tends to be asking why something happened. Feynman said you can always ask why once more – why, why, why, why... So, a scientific explanation in terms of physics is creating some sort of model, often a combination of a physical intuitive nature and a description or a model that when you combine together you can predict what will happen before you do it. In terms of biology and progressively softer sciences it becomes less and less why and more what. Chemistry is in the middle - if you do this what will happen? And then you go to physics and ask why did that happen? … So, it varies and depends on which field.

A distinct observation made in the case of scientists, as compared to other participants, was that the most important feature or theme of a scientific explanation for scientists was that it includes elemental features that are logically coherent. Therefore, a scientific explanation is seen an investigative or experimental process where elements used to construct the explanation are empirically testable and are logically connected with observations and scientific laws. What is more, it is important to note that scientist participants, unlike teachers and students, did not consider mathematical equations or mathematical proof as an essential feature to scientific explanation. Table 4.11 sums up the generated features from scientist transcripts with accompanying frequencies regarding what scientist participants consider being an explanation.

Table 4.11

Scientists' Features of a Scientific Explanation

Feature	Illustrative excerpt	*f*
A scientific explanation is based on facts (i.e. various forms of data or evidence, connections observations with facts and laws)	Sophia: It [A scientific explanation] is based off of facts that we know and observables that we observed. We can correlate the facts and laws to explain our observables.	9
A scientific explanation is logically coherent (i.e., has a logical flow and internal consistency of scientific ideas)	Selena: You have some number of ingredients and you want to put them together, and you tell a self-contained story that should flow.	7
A scientific explanation is empirically testable, has predictive power (i.e., able to predict some phenomenon before it occurs), verifiable, and repeatable.	Stanley: It [A scientific explanation] should provide a consistent framework for understanding and predicting whatever it is trying to explain.	6
A scientific explanation should be statistically or probabilistically true (i.e., it should be true a reasonable number of times)	Samantha: It [A scientific explanation] has to be provable to some degree. You can have a theory model that you think explains the world but it is not enough, that's not an explanation. You have to confirm your theory with experimental data and only then it becomes an explanation that you could rely on and you could share with other people - that is to some reasonable extent by analyzing statistical probability.	6
A scientific explanation must include all relevant observations, facts, and laws (and dismiss all irrelevant ones)	Stanley: For every scientific explanation we try to get to understand a certain number of relevant scientific concepts and these should be introduced. Everything has to be considered - even if they have to be discarded.	5

(Table continues)

Table 4.11 (Continued)

Feature	Illustrative excerpt	*f*
A scientific explanation follows the scientific method.	Stella: It [A scientific explanation] is basically communicating how something works based on a process that follows the scientific method.	4
A scientific explanation provides understanding (i.e., it makes sense of the natural world).	Sam: A scientific explanation is less about how something happens but more of a way of looking at things and figuring out what are the relevant things to look at... So it is more about be able to understand that situation.	4
A scientific explanation is field-dependent (i.e., explanations in hard sciences are different than those in softer sciences)	Stefan: In the physical science, not in the social sciences or psychology, it [a scientific explanation] is something that is based on observable, repeatable observations and descriptions.	2
A scientific explanation is clear and comprehensible.	Sylvia: It [a scientific explanation] should be clear using a language that is not specific to a scientific discipline or if it is it should be explained clearly in a way that anyone can understand. It should be at a level that anyone could understand regardless of scientific expertise.	2
A scientific explanation is context-dependent (i.e., it depends on the learner's prior knowledge)	Sylvia: "A scientific explanation is communicating science in a way that makes sense to the individual you are trying to explain it to, which will depend on their level of knowledge and how much they know about science."	2
A scientific explanation is objective or always true (i.e., does not depend on opinion, is not subjective)	Stefan: A scientific explanation is any explanation that the answer depends not on subjective thought or opinion but rather observable and repeatable facts.	1

Commonalities and Differences of Perceptions. All three participant groups mainly agreed that a scientific explanation relies of scientific knowledge, observations, inferences, and scientific laws and theories. However, the relationship and the role that each of the aforementioned features play in a scientific explanation were somewhat different across the groups. Nonetheless, all groups did emphasize that the major role of scientific knowledge (referred to in NOSE framework as pieces of knowledge) and the role of inferences were to provide support for the observations made.

Students and teachers also considered that a scientific explanation needed to be unbiased. It was noted that scientific explanation must not leave much up to someone's assumptions, must not be based on a person's opinion, and must not leave room for misinterpretations. What is more, students and teachers also believed that scientific explanation must not be subjective and must be *always* true or correct. However, teachers added that scientific explanation must depend on the explainer's (or the recipient's) prior knowledge (i.e., an explanation is context-dependent).

While students based their perception of scientific explanation mainly on observations, teachers and scientists focused on how explanation relies on scientific knowledge, inferences, principles, and laws in relation with observations. In addition, teachers and scientists further differed from students in their consideration of the quality of elemental features needed in the construction of a scientific explanation (e.g., testability, reproducibility). However, unlike both teachers and students who did not consider logical coherence as a major feature of a scientific explanation, scientists emphasized flow and internal consistency (i.e., logical coherence) as an important feature for developing an explanation. Figure 4.30 provides a graphical comparison of the views of all three participant groups regarding the nature of explanation.

Participants Criteria for Assessing Scientific Explanations. This section presents the criteria for assessing scientific explanations when participants were asked to examine and assess each other's' explanations about the four scientific phenomena. Participants were also asked to list, in their own words, the criteria they used to assess the adequacy of a scientific explanation. The criteria had both major (articulated by, at least, 50% of participants in a group) and minor (articulated by less than 50% of participants in a group) themes. The following major criteria were common to, at least, two participant groups: correctness or accuracy, use of scientific information and evidence to support observations, logical flow, sufficient of content knowledge included, learner appropriateness, and thought alignment.

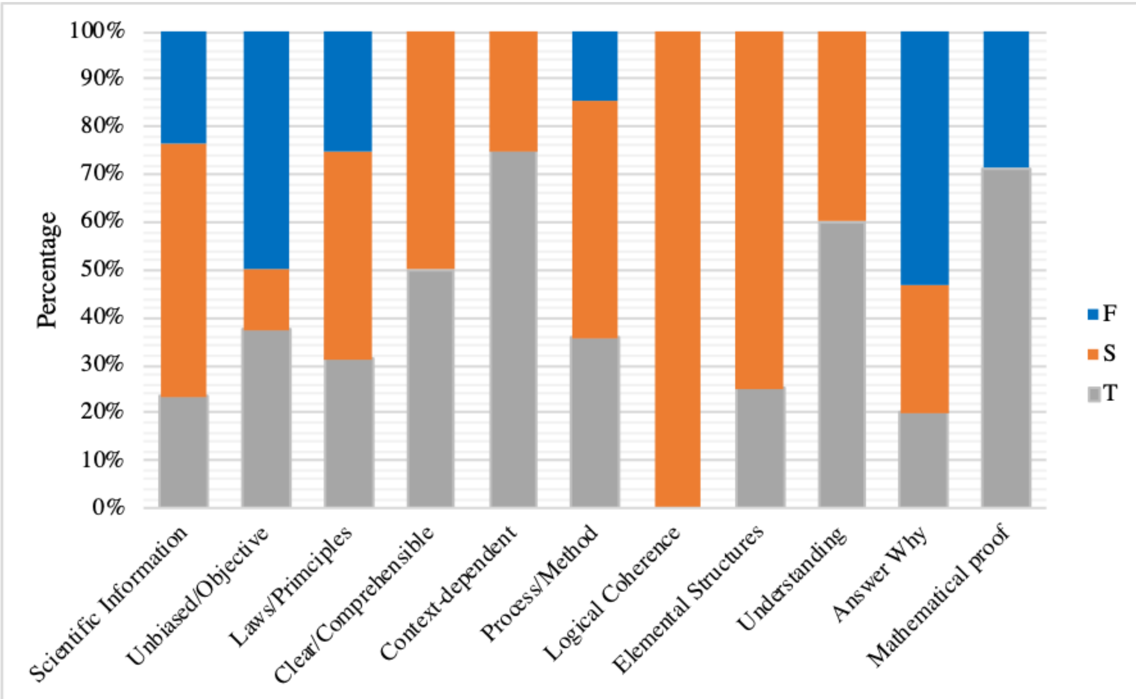


Figure 4.30. Participants' features of scientific explanation. Features are shown where F represents freshman students, T-teacher and S-scientist. The features are based on percentage frequencies. Note that Clear/Comprehensible and Context-Dependent on the X-axis are two features that show commonalities between only two participant groups.

Common minor themes were related to the use of all and only relevant scientific information, simplicity of explanation, depth of explanation (i.e., further explanations of ideas), clarity, comprehensibility (i.e., sense making), sufficient description of the phenomenon, and connectedness to real life. The following sections discuss both major and minor criteria across and within all participant groups.

Students. When asked to evaluate peer and other participant group explanations of the different scientific phenomena, the majority of students (6 of 10) considered comprehensibility, identified mainly as making sense, of an explanation as a major criterion of a good scientific explanation. It did not mean that scientific ideas were clearly connected. It merely implied that, what the explainer was stating could be understood by the participant (in this case the student participant). Finn, for example, while assessing a teacher's explanation of the dancing raisins scenario, explicitly stated:

> This does not make sense to me. I don't think it makes sense that the bubbles would be able to enter the raisins… I don't know… He… or she does not explain how… I think this is why I didn't choose this as the best explanation… It just does not make sense to me.

Another major criterion of a good scientific explanation was related to the content knowledge included in a given explanation. In particular, 8 of 10 students considered including prior content knowledge to be a feature of a good scientific explanation. For example, Franco noted:

> I think they [the scientist] know what they are talking about here. This would be good in an explanation… to include your background knowledge, so you can explain it well. They mention the ideal gas law and they say what it is… so that's definitely a good thing.

Seven of 10 students considered correctness or accuracy a criterion of a good scientific explanation. Students judged the validity of an explanation based on what they considered to be correct. For example, when asked about the criteria of a good scientific explanation, Faith noted: " [A good scientific explanation] has to be accurate. It has to make sense according to scientific laws.".

In addition, most students (6 of 10) tended to look at how well the explainer described the phenomenon-to-be-explained. For example, Florence considered an explanation to be good because “it included good and sufficient descriptions of the phenomenon.” She further noted that “without knowing what is going on, we cannot understand why it happened.”

Logical flow was identified by 5 of 10 students as another criterion of a good scientific explanation. It did not mean that there was only one order by which a scientific explanation had to follow. It meant, however, that what the explainer was stating had to be of a logical order that was clearly shown in the explanation. Fidel, for example, stated: “I don’t think necessarily there is a right or wrong order, but it is important to note which order it is in - whether this explanation first, or the observation first.”

Interestingly, student participants considered the various ways of explanation (noted as thought alignment) about the phenomenon to be a criterion of a good scientific explanation. Thought alignment mainly emphasized whether or not the explainer exhibited a similar way of thinking about the phenomenon-to-be-explained as the participant did (5 of 10), and whether or not the explainer answered a particular *why*-question using somewhat similar facts, laws and principles. This criterion was dependent, from the students’ perspective, on the level or prior knowledge of the explainer, which was important in determining the strength and weakness of an explanation. For participant students, that depended on the background of the explainer (e.g., if they had a clear scientific background about the phenomenon). For example, Fidel shared how he assessed the various scientific explanations, and he noted:

> The way I thought which explanation was good was whether they thought of it the same way I did. And this explanation is not good because they say it’s convection. This is not why the raisins go up. They go up because of density. It has nothing to do with convection…. I think this is what I said. Density… not convection.

Finally, clarity was an aspect of a good scientific explanation reported by half of the students. For instance, Farrah noted that a good scientific explanation has "to be straight to the point. It has to answer the question right away and then has to add clear details as to why it is that."

Minor criteria generated by student participant responses were simplicity, connectedness to real life, use of scientific information to support observations, relevance of scientific information, depth of explanation, and learner appropriateness. Table 4.12 shows an illustrative quote for each criterion. Simplicity was important to 4 of 10 students when assessing an explanation. In addition, connectedness to real life was another aspect to a good scientific explanation, as seen by four students. Three students emphasized the use of scientific information to support observations, as well as, the use of only relevant information as important criteria. For these students, a good scientific explanation was one that included only scientific information that supported their observations and were related to the phenomenon-to-explained. Depth of explanation was a criterion mentioned by three participant students. Depth implied whether or not the explainer provided further explanations to every idea or fact they included. Finally, two students believed that the validity of an explanation depends on the learner (i.e., the recipient of the explanation).

In conclusion, when assessing explanations made by their peers and other participants, students tended to use criteria emphasizing attributes to *both* the explainer and explanation. First, major features of the explainer were those that required them to adequately describe the phenomenon before explaining it, explain in a way that is accessible to the participant, and demonstrate correct personal content knowledge of the phenomenon-to-be-explained. Minor features of the explainer included using examples from everyday life to clarify less familiar phenomena, explain all relevant ideas associated with the phenomenon (i.e., explain in depth), and include only information related to the phenomenon-to-be-explained.

Table 4.12

Minor Criteria of Explanation Articulated by Students

Minor criterion	Illustrative quote	*f*
Simplicity	Filip: I would look at an explanation that uses one principle or the least number of principles. It has to be simple in that sense.	4
Connectedness to real life	Fiona: I like this explanation because it has an example. I understand this now because it is true… when we put floaties in the pool we stay up… we float. So examples are good in an explanation. They make it better…. I have to say examples are one of the things that make an explanation…. good, or better.	4
Use of scientific information to support observations	Faith: I think what makes this a good one [explanation] is that they explain that these bubbles are air bubbles and then they say that air is less dense than pop. And this is good because we can see the bubbles on the raisins.	3
Relevance of Scientific in formation	Flynn: A good explanation has to consider all the factors. It has to incorporate the core idea of that specific experiment.	3
Depth of explanation	Fredrick: I don't think this is a good explanation. They [a student] say, "this has to have been a product or a result of the combustion of the candle" so the pressure decreases. So what? They do not explain what is it the causes the pressure to decrease. I think it is important to keep asking yourself why.	3
Learner appropriate	Faith: A good explanation depends on who is explaining or reading the explanation. It all depends on that.	2

Second, major features of the explanation included the explanation logical flow among its elements, and whether or not the explanation is clear and comprehensible (i.e., made sense to the assessor). Minor features of the explanation included the presence of relevant information supporting the observations made, its simplicity, and whether or not it is learner appropriate.

Teachers. All teacher participants considered the structure of explanation (also identified by teachers as the need for a good explanation to have a clear structure) as the most important criterion in evaluating a scientific explanation. Four of the 10 participant teachers mentioned that in assessing an explanation they look for a claim-evidence-reasoning structure. For example, Tucker noted:

> What I look for is a claim - what is actually happening, and your reasoning - here's why this is happening. And if you provide a reason that you can support by evidence or some sort of a test then that's a good explanation.

The remaining teachers emphasized other structures: three considered following the scientific method as something they would look for in assessing a scientific explanation, while the remaining three teachers considered a general structure (i.e., the explanation should have a certain structure). For example, when sharing the criteria that he would use as a teacher to assess the goodness of a scientific explanation, Tod noted:

> A good scientific explanation is one with concepts that contain some sort of experimental basis within the scientific method along with some actual scientific concepts that has to be named …. It should show the scientific method. How you made your observations, how you collected your data or added your evidence. And what conclusion you reached … and how you reached that … this last part would be where the concepts would be added.

In addition, all teachers considered a good scientific explanation as one that is appropriate to the learner receiving the explanation. When asked about the criteria they looked for in assessing the goodness of a scientific explanation, all teachers thought of this task as assessing their students' explanations. Hence, in listing these criteria, all 10 teachers were explicit in saying that what they mainly look for is whether or not the explanation is adequate to the student level, whether the student used their prior content knowledge or what they have learned in class, and whether or not they could teach it to their peers. Some teachers further added learner-dependent as a criteria of a good scientific explanation they, as teachers, would provide for their students in a classroom setting. For example, Thomas noted that “an explanation needs to be appropriate to the learner and not simply provide the most exact understanding of the whole thing as, say, a science researcher, would explain it [the phenomenon]”.

Unlike students, the majority of teachers (9 of 10) considered being correct an essential criterion of a good scientific explanation. Finally, half of the teachers considered relevance of scientific information and laws an essential criterion of a good scientific explanation. For example, Tyson noted:

> A good scientific explanation is one that includes everything… it should not leave facts that are important to the explanation… I guess this depends on prior knowledge as I said earlier… but yeah, all facts… all laws too… when you talk about why food coloring spreads faster in warm water than in cold water, you should mention temperature, but you should also mention the law… everything else. That’s another thing I would look for. But that depends on who is giving an explanation.

Most teachers (9 of 10) also considered that the explainer needed to demonstrate sufficient content knowledge to help strengthen their explanations. Tucker, in his evaluation of a student transcript noted “they are trying to explain why they see two pennies, but they don’t mention refraction of light or why light refracts. They only talk about air and water. They don’t have the correct knowledge of refraction of light.”

Examples and connectedness to real life was another criterion of a good scientific explanation for 7 of 10 teachers. When asked about the criteria of a good scientific explanation, Tracy noted that “a good explanation has aspects of what you see and how it relates to science and everyday life. If you cannot link it to your life, then you probably don’t understand it.” Six teachers tended to look at how well scientific information supported observations in a given explanation. Tyson, in his evaluation of a teacher transcript noted "I like this explanation because every observation is backed up by data … by evidence. They probably did not have a chance to actually collect data … I know I did not … but they still support their observations with scientific concepts." Interestingly, half the teachers considered a good explanation as that which is similar to their way of thinking about the phenomenon at hand, although three teachers mentioned that they needed to collect more data or conduct more research to make sure that their way of thinking was accurate. For example, Tarra thought that the scientist's explanation was the best one because "it is how I would explain it to my students. It cuts to the chase. I like it because I think it is correct."

Minor criteria generated from teacher assessment also emphasized criteria about relevance of scientific information, clarity, comprehensibility, simplicity, depth, and descriptions of phenomena. Four teachers considered that a good scientific explanation is one that is clear, simple and includes all and only relevant information. In addition, four noted that a good explanation should be understood by the recipient (i.e., its comprehensibility). Similar to students, another minor criteria (3 of 10) was related to the depth of the explanation (i.e., whether or not all why-related questions are answered). Finally, like students, two teachers believed that in order to explain why something is happening, one needs to describe it first. Table 4.13 presents quotes that are illustrative of minor criteria identified among teacher participants.

In conclusion, when evaluating explanations made by their peers and other participants, teachers tended to use assessment criteria emphasizing features of *both* the explainer and the explanation. Major criteria of the explainer were those that required the explainer to clearly and correctly state their answers of the *why*-question asked, demonstrate sufficient content knowledge of the phenomenon-to-be-explained, use examples from everyday life and connect the phenomenon at hand to real life, and explain the phenomenon in a way similar to that of the participant/assessor. Minor features of the explainer were related to their ability to explain all related observations and ideas, describe the phenomenon at hand, and include all and only relevant facts associated with the phenomenon-to-be-explained. Major criteria of the explanation included its structure (i.e., logical flow), whether it is learner-appropriate, and whether the scientific information included support the observations made. Minor features of an explanation included its clarity, simplicity, and comprehensibility (or sense making).

Table 4.13

Minor Criteria of Explanation Articulated by Teachers

Minor criterion	Illustrative quote	*f*
Use of relevant of scientific information	Tanya: I think something to look at in an explanation is to check if the explanation has all the facts that should be there. A lot of the time students add everything they know… irrelevant things. And they hope that one thing is.	4
Comprehensibility	Thomas: I also tell them [my students] that it [the explanation] needs to make sense. It has to be understandable, not just by them, but by anyone who reads it.	4
Clarity	Tammy: If an explanation is not clear then I won't be able to even grade it. It first has to be clear… meaning that you want to be very clear about what you are saying… a lot of the time students just go around …. You should say what you want to say right away… I am also thinking about what kinds of words you are using… they have to be clear… not nonscientific... but then you should define them if you say pressure for example. What is pressure?	4
Simplicity	Tina: A good one [scientific explanation] must be able to explain in a simple way… You need to explain the natural phenomenon in order to make sense of the world… if it is too hard, you can't make sense of it… it has to be easy… or simple.	4
Depth of explanation	Tucker: I don't like this explanation [points to a student's transcript]. It leaves me asking a lot more why questions. It does not explain everything. A good explanation should answer all the questions.	3
Sufficient descriptions of phenomenon	Trevor: It [A good scientific explanation] is one that contains true scientific principles that can be supported and they should match your observations... And it is important to make a good observations to begin with... It is important to include all your observations.	2

Scientists. All participant scientists considered the existence of all and only information related to the phenomenon-to-be-explained as the most important criterion when evaluating scientific explanations. What is more, scientists also considered dismissing information that is not relevant to the phenomenon. For example, Sara commented, "it is important to leave out things that are not helpful and not just include things that are." In addition, Saul noted that for him a valid scientific explanation was one that "includes only the laws and facts that matter…that depends on prior knowledge." In addition, all scientists emphasized that a good scientific explanation is one that includes scientific information that is aligned with the observations made. Scientists also considered that scientific information (i.e. scientific prior knowledge) and observations be supported by scientific laws. For example, when evaluating a student's transcript, Samantha noted: "it seems like a reasonable explanation because you can clearly break it into observations then conjectures … or facts that in turn invoke some axioms or laws."

What is more, the majority of scientists (9 of 10) considered correctness or accuracy to be an essential criterion of a good scientific explanation. Scientists believed that a good scientific explanation should include correct and "factual information." For example, Stanley noted that a good scientific explanation "should be correct, repeatable and consistent. I think that it is more important to be right than to be clear." In addition, 7 of 10 scientists considered logical flow of an explanation was essential for constructing a good scientific explanation. These scientists emphasized that a good scientific explanation should be organized in a way that showed consistency among the ideas presented, such that the overall explanation was comprehensible. For example, Sam noted:

> A good scientific explanation is one that can bring it home in a way that satisfies us in a story like manner based on observations and descriptions that leads to factual statements that can be made which then can lead to lawlike statements.

Finally, similar to students and teachers, demonstrating sufficient content knowledge was considered to be an important criterion of a scientific explanation among six scientists. For example, in evaluating a teacher's transcript, Selena noted: "I feel like they need to add more words here… more details. This feels like a draft of an explanation for lack of a better word." Finally, six scientists believed a good scientific explanation should be learner-appropriate. For example, Sylvia considered a good scientific explanation to be "clear and uses language that is not specific to a scientific discipline or if it is it should be explained clearly in a way that anyone can understand."

There were minor criteria noted by 4 of 10 scientists and less. These were comprehensibility, depth of explanation, clarity, connectedness to real life, and simplicity. In particular, four scientists considered a good scientific explanation is one that is comprehensible (i.e., is understandable) and explains all why-related questions. In addition, three scientists emphasized clarity and the use of everyday examples as attributes of a good scientific explanation. These scientists believed that including examples from everyday life makes a scientific explanation better. That meant that a good scientific explanation is one that is relevant to the explainer's and the recipients' everyday lives. Finally, only two scientists believed that a good scientific explanation should be simple. That meant that a good scientific explanation should not include confusing or complicated ideas and is understandable by non-scientists as well as scientists. Table 4.14 presents sample quotes illustrative of minor themes invoked by scientists.

In conclusion, while evaluating explanations made by their peers and other participants, scientists tended to use criteria that emphasized features of both the explainer and explanation. Major criteria of the explainer were using correct scientific concepts, demonstrating sufficient content knowledge, and including only relevant information (and dismissing irrelevant information). Minor criteria of the explainer included using examples from everyday life to make sense of the phenomenon at hand and explaining in-depth all related observations and ideas. On the other hand, major criteria of the explanation included its logical flow and connectedness of its elements, the alignment between its elements, and its learner appropriateness. Finally, minor criteria of the explanation included its clarity, comprehensibility and simplicity.

Table 4.14

Minor Criteria of Explanation Articulated by Scientists

Minor criterion	Illustrative quote	*f*
Comprehensibility	Selena: One of the most important criteria for me is that what they are explaining simply makes sense. Does the reader understand it? If yes, then it is a good explanation.	4
Depth of explanation	Stefan: This explanation does not feel good. I would want to know more. For example, why does the pressure decrease inside. They make it sound like for some unknown reason, pressure inside decreased.	4
Clarity	Saul: It [The scientific explanation] has to be clear…. you should say what you want to say clearly.	3
Connectedness to real life	Sara: I also think that the third element is a way to effectively connect with your listener through the use of examples. It is a way to connect with people outside of science.	3
Simplicity	Stella: One criterion is being simple so that anyone could understand it regardless of scientific expertise. If you cannot explain something to a 10-year-old then you do not really understand it… I think Einstein said that. And I think that is an important criterion.	2

Common criteria within and across groups. Both major and minor themes were common among the criteria articulated by the three participant groups. Major themes focused on the attributes of both the explainer and the explanation. These were correctness or accuracy, logical flow, and sufficient of content knowledge presented. Simplicity and depth of explanations were common minor criteria across the three groups. Figure 4.29 presents the major criteria that were common across the three participant groups (common to at least two groups). Among the major criteria presented in Figure 4.31, all participants considered the logical flow of explanation. That is, all participants believed that a strong scientific explanation was one that has its elements connected logically in a coherent way. In addition, all participants tended to look for canonical correctness and sufficient content knowledge while assessing scientific explanations constructed by their peers or participants from other groups.

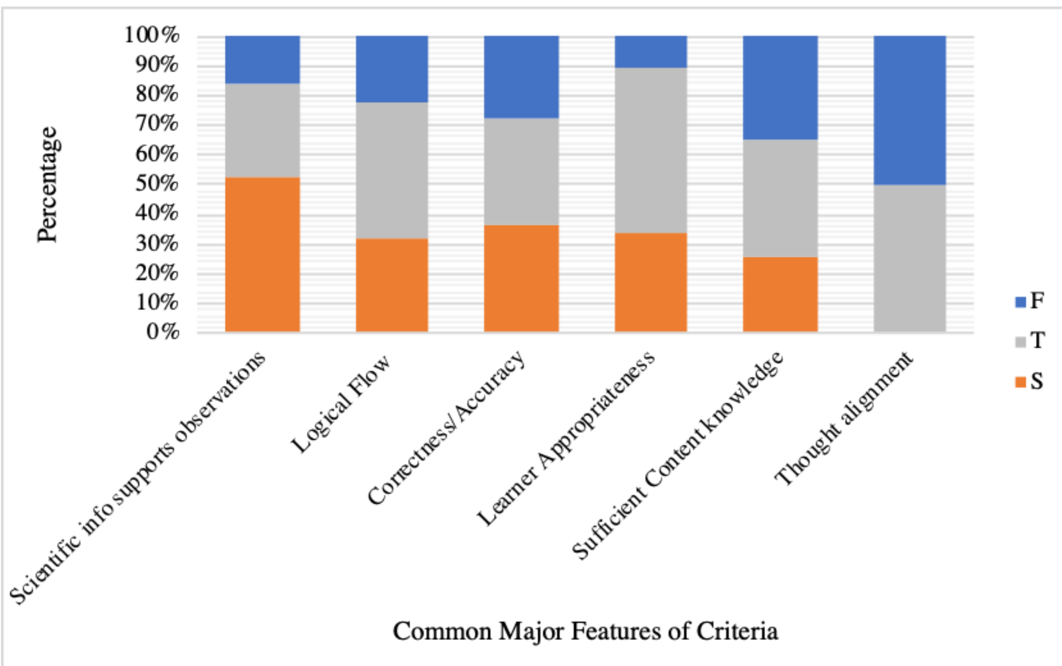


Figure 4.31. Major common criteria. The major common criteria (stated by 50% or more in each group) used across participant groups represented by F-freshman students, T-teachers, and S-scientists. The features are based on percentage frequencies. Note that Thought Alignment criterion on the X-axis shows commonalities between only two participant groups.

Students and teachers (though not mentioned by scientists) emphasized thought alignment that needed to be similar to that of the assessor. In their articulation about thought alignment, students and teachers indicated that a good explanation was one based on the fact that the explainer used similar scientific information, laws and concepts as they did when they provided their explanations. In addition, significantly more teachers (all teachers) than students and scientists considered a good explanation to be one that is learner appropriate (i.e., pedagogically accessible). Finally, significantly more scientists (all scientists) than students and teachers believed that a good scientific explanation was one that connects scientific concepts, observations and laws together.

Common Minor Criteria within and Across Groups. Among the common minor criteria in at least two of the three groups emphasized attributes of the explainer, such as, the depth of explanation that needed to provide answers to all relevant *why*-questions, the use of all and only relevant information, the use of examples to connect the phenomenon-to-be-explained to everyday life, and the importance of descriptions in an explanation. Finally, comprehensibility was considered to be a minor criterion in at least two of the three participant groups.

While all scientists believed that including all and only relevant scientific information as an essential criterion of a good scientific explanation, less than half of the students and teachers considered this to be a factor while assessing explanations. On the other hand, the majority of teachers emphasized that connecting an explanation to real life is a strength of an explanation, as opposed to less than half of the students and scientists emphasizing connectedness to real life as an aspect of a good scientific explanation. Figure 4.32 also shows minor criteria that were common to at least two of the three groups.

Perceptions of the "Goodness" of the Explanations Made by Peers and Others

As they assessed explanations made by their peers and by members of the other two groups, participants often made overall judgements about different aspects related to the “goodness” of these explanations. The reader is reminded that due to the relatively small sample size of the study, in the second interview, explanations of the same scenario were not presented to the same participant more than once. For example, each participant scientist examined a randomly selected explanation of a given scenario from among explanations generated by the other nine scientists, one explanation of a different scenario selected from those generated by the 10 participant teachers, and one explanation randomly selected from the 10 participant student explanations (also of a different scenario).

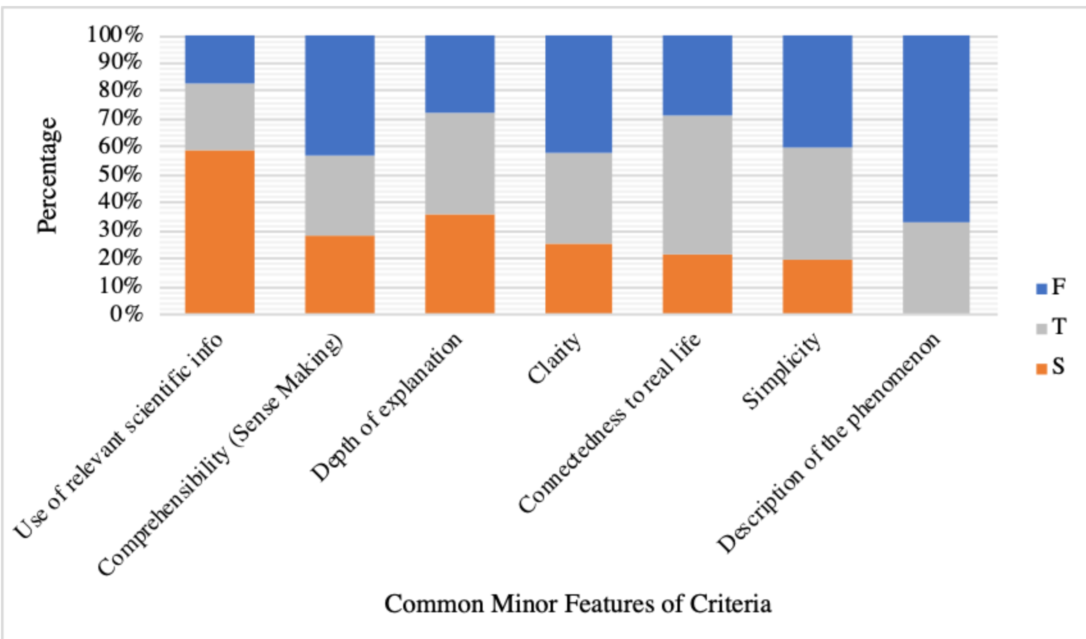


Figure 4.32. Minor common criteria. The Minor common criteria (stated by less than 50% of participants in each group) used across participant groups represented by F-freshman students, T-teachers, and S-scientists. The features are based on percentage frequencies. Note that Description of the Phenomenon criterion on the X-axis shows commonalities between only two participant groups.

Participants then assessed and provided feedback on final explanations generated during the first phase by one participant from their own group, as well as one form each of the other two groups. More specifically, participants were asked if these explanations were complete/incomplete (completeness), and if they were adequate, partially adequate, or inadequate (adequacy). Participants were also asked to indicate the "best" explanation in the set of explanations they were evaluating. Hence, the result was a total of 30 assessments of explanations generated by each group, that is, there was a total of 30 assessments each of scientist explanations, teacher explanations, and student explanations. Table 4.15 shows the random assignment of explanations of all participants during the second interview.

Tables 4.16 and 4.17 summarize the results of these assessments. Overall, participants judged each transcript as providing a complete and good (adequate) explanation (symbolized by "CA" in Tables 4.16 and 4.17); a partially adequate or incomplete explanation (PA) where a participant was judged to have started constructing an explanation but failed, for instance, to include all scientific information necessary to explain the phenomenon; an inadequate explanation (IA) where a participant was judged to have constructed an explanation using inaccurate facts, laws, etc., to explain the phenomenon; or a non-explanation (NE) because a participant's transcript, according to the assessor, was limited to, for instance, describing rather than explaining the phenomenon. The majority of participants also designated one of the three transcripts they examined as providing the "best" explanation in the set (labeled with a cross in Table 4.16).

An examination of Table 4.17 and Figure 4.33 shows that, overall, scientists did much "better" than the participant students and teachers. From the perspective of participants, a majority of scientist transcripts (63%) were judged to present complete and adequate (or good) explanations, which was much larger than the percentage of teacher (40%) and student (13%) transcripts deemed to present complete and adequate (or good) explanations. Similarly, more than half (53%) of the student transcripts and 30% of teacher transcripts were judged as either not making an explanation or constructing an inadequate explanation, compared to only 10% of the scientist transcripts.

Table 4.15

Interview II Random Assignment of Explanations Across the Three Groups

Participant	Peer	Other Group	Other Group
S01	S03 Food Coloring*	F06 Penny	T02 Raisins
S02	S04 Food Coloring*	F07 Penny	T10 Candle
S03	S09 Food Coloring	F08 Raisins	T10 Penny
S04	S02 Candle	F05 Food Coloring	T07 Penny
S05	S07 Candle*	F05 Raisins	T05 Penny
S06	S02 Penny	F08 Food Coloring*	T08 Raisins
S07	S08 Candle	F10 Raisins	T09 Penny*
S08	S10 Penny*	F06 Food Coloring	T01 Candle*
S09	S06 Food Coloring*	T02 Candle	F03 Penny
S10	S08 Food Coloring	F08 Penny	T09 Raisins*
F01	F05 Candle	T03 Raisins*	S10 Food Coloring*
F02	F04 Raisins	T08 Penny	S02 Food Coloring*
F03	F04 Candle	T01 Food Coloring*	S06 Penny*
F04	F07 Candle	T02 Penny	S01 Raisins*
F05	F09 Candle	T05 Food Coloring*	S05 Raisins
F06	F01 Raisins*	T04 Food Coloring	S08 Penny*
F07	F01 Candle	T01 Penny*	S06 Raisins
F08	F07 Raisins*	T07 Food Coloring	S04 Penny*
F09	F06 Raisins*	T05 Candle	S03 Penny
F10	F04 Penny	T01 Raisins*	S01 Food Coloring*
T01	T03 Candle	F01 Food Coloring	S01 Penny*
T02	T04 Candle*	F02 Penny	S03 Raisins*
T03	T06 Candle*	F02 Food Coloring	S07 Raisins*
T04	T03 Food Coloring*	F10 Penny	S09 Candle
T05	T06 Penny	F10 Food Coloring	S04 Candle*
T06	T09 Candle	F02 Raisins	S05 Penny*
T07	T10 Food Coloring	F01 Candle	S09 Penny
T08	T07 Candle	F03 Penny	S10 Raisins*
T09	T06 Raisins*	F03 Candle	S07 Penny*
T10	T04 Raisins	F03 Food Coloring	S05 Candle

*Denotes final explanations that were judged by other participants to present a complete and adequate (or good explanation).

About 33% of participant students and 30% of teacher transcripts were also judged to have presented only partially adequate (or incomplete) explanations compared to 27% of the scientists. Finally, more than half (53%) of the scientist transcripts were judged to present the "best" explanation within their particular set as opposed to 23% of teacher transcripts and only 10% of student transcripts.

Table 4.16 also shows that judgements were not necessarily differential by group. In other words, the results are not skewed by the judgements of one group of participants. For instance, only half of the students judged scientist transcripts to provide the "best" scientific explanation in a particular set. This is comparable to six teachers and five scientists who made a similar judgement. In addition, 7 of 10 students judged scientist transcripts to provide complete and adequate (or good) explanations. This is also comparable to seven teachers and five scientists who made a similar judgement. Similarly, 5 of 10 students judged teacher transcripts to provide complete and adequate (or good) explanations as compared to four teachers and three scientists.

Additionally, only three students, no teachers, and one scientist thought that students made complete and adequate (or good) explanations. Thus, it could be concluded that, from the perspective of all participants, scientists seemed—by far, to make the better explanations. In other words, when building their explanations, scientists seemed to have addressed a larger number of criteria that were deemed by participants to be essential or important to making good explanations.

> In the present study, students, teachers and scientists all agreed in their judgement about who presented "good" explanations, namely, scientists. This finding is not surprising. It is aligned with current research that considers the practice of explanation by scientists to be the framework or the standard. Yet, this finding provides robust support for the validity of the NOSE as a framework that would eventually enable the assessment of learner-generated explanations.

However, further examination of the results reveals an interesting finding when results from Tables 4.16 and 4.17 are compared with data from Table 4.15. In particular, Table 4.15 presents the random assignment of the "sets" assessed by all by all participants.

Table 4.16

Summary of the Assessments by Participants of the Explanations of Peers and Other Participants

	Students			Teachers			Scientists		
Assessor	F	T	S	F	T	S	F	T	S
1	NE	CA	[a]CA	PA	PA	[a]CA	IA	IA	[a]CA
2	IA	PA	[a]CA	IA	CA	[a]CA	PA	PA	[a]CA
3	NE	[a]CA	CA	NE	CA	[a]CA	PA	IA	PA
4	PA	IA	[a]CA	IA	[a]CA	NE	PA	PA	PA
5	IA	[a]CA	PA	NE	NE	[a]CA	IA	NE	[a]CA
6	[a]CA	PA	CA	IA	IA	[a]CA	[a]CA	PA	IA
7	PA	[a]CA	PA	PA	PA	PA	PA	[a]CA	IA
8	CA	IA	[a]CA	NE	IA	[a]CA	PA	CA	[a]CA
9	[a]CA	PA	PA	PA	[a]CA	CA	NE	PA	[a]CA
10	NE	CA	[a]CA	IA	IA	PA	IA	[a]CA	PA

Note. CA= Complete Adequate, PA = Partially Adequate (or incomplete), IA = Inadequate, and NE = Non-Explanation.
Assessors are shown where F represents freshman student, T-teacher and S-scientist.
[a] An indicator of those explanations considered to be best explanations in a set. In Table 4.16, it indicates the total amount of best explanations within each set.

Table 4.17

Summary Totals of the Assessments by Participants of the Explanations of Peers and Other Participants

	a		CA		PA		IA		NE	
Groups	n	%	n	%	n	%	n	%	n	%
Students	3	10	4	13	10	33	9	30	7	23
Teachers	7	23	12	40	9	30	7	23	2	7
Scientists	16	53	19	63	8	27	2	7	1	3

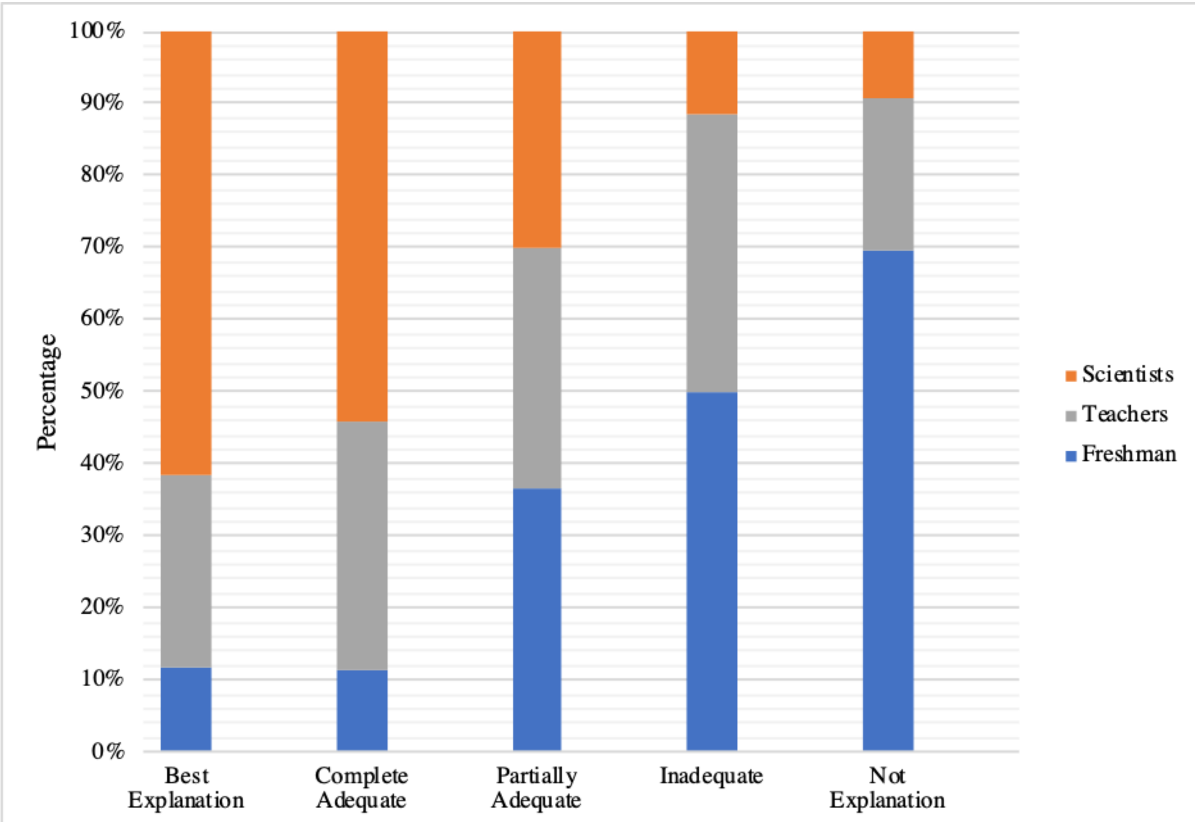


Figure 4.33. Summary of the Assessments by Participants of the Explanations of Peers and Other Participants

For example, Table 4.15 shows that scientist S01 assessed: (1) scientist S03 final explanation of the Food Coloring scenario, (2) freshman student F06 final explanation of the Penny in a tub of water scenario, and (3) teacher T02 final explanation of the Dancing Raisins scenario. Table 4.18 shows that while the judgements are not necessarily differential by group, they are differential by scenario. In other words, these results show that participant groups tended to *prefer* explanations from certain phenomena over others. Hence, this preference might have played a role in their judgements of a complete and adequate explanation. For instance, only three teachers and two scientists judged explanation transcripts of the Burning Candle scenario (i.e.. the water rising phenomenon) as complete and adequate explanations; while none of the students did (even though 6 of 10 students were randomly assigned transcripts of the Burning Candle scenario). On the other hand, more students (6 of 10) and teachers (4 of 10) judged explanation transcripts of the Dancing Raisins scenario as complete and adequate explanations; while only one scientist did (even though 6 of 10 scientists were randomly assigned transcripts of the Dancing Raisins scenario). Additionally, five students and four scientists judged explanation transcripts of the Food Coloring scenario as complete and adequate explanations; while only one teacher did (even though 6 of 10 teachers were randomly assigned transcripts of the Food Coloring scenario). Finally, four students, three teachers, and two scientists judged explanation transcripts of the Penny in Water scenario as complete and adequate explanations.

Hence, out of the 35 complete and adequate explanations as judged by all participants, 11 were explanation transcripts of the dancing raisins scenario, 10 of the food coloring, nine of the penny in water, and only five were of the burning candle scenario. This finding could be related to the complexity of the Burning Candle phenomenon, and the fact that many participants—especially among students and teachers—were not able to generate a complete and/or adequate explanation of this phenomenon. It could also be due the fact that students could not make sense of the phenomenon itself or could not, simply, comprehend the explanation they were judging. In fact, the burning of a candle inside an inverted jar partially immersed in water is an experiment that was done over 2,000 years ago (Vera, Rivera, Nunez, 2011). To this day there is a common misconception that many people and even science textbooks use to explain the water rising due to oxygen consumption in the air – a result that was also found in the present study. In addition, this finding could be related to the pragmatic nature of scientific explanation. In other words, from the perspective of participants, a complete and adequate (or good) explanation is related to the assessor's level, their prior content knowledge, and their confidence of the scientific phenomenon related to this demonstration.

Table 4.18

Summary Totals of the "Complete and Adequate (CA)" Explanations as Judged by Participants Across the Four Scientific Phenomena

	Candle	Raisins	Food Coloring	Penny
Groups	n	n	n	n
Students	--	6	5	4
Teachers	3	4	1	3
Scientists	2	1	4	2
Total	5	11	10	9

Participants Perceptions in Relation to NOSE Structural Elements. This final section addresses the fourth and last research question: to what extent do participants' perceptions of valid scientific explanations draw upon elements that are characteristic of the NOSE framework? This section starts with a brief overview of the NOSE characteristics and structural elements of scientific explanation. This is followed by a discussion of participants' perceptions of explanation as they relate to the NOSE framework.

The NOSE's view of explanation. According to the NOSE framework, scientific explanation is pragmatic. In other words, NOSE emphasizes that explanation-statement(s) necessitate(s) a reference to an explainer that is appropriate to a certain context. Stemming from philosophical models of scientific explanations, NOSE framework argues that non-pragmatic criteria such as derivability from laws and causation, are not sufficient to judge the "goodness" of an explanation. An explanation is rather judged as adequate not because it answered a causal question in a unifying lawlike manner, but because it did so at a level that is adequate and understandable to the explainer and/or the recipient of the explanation. Second, NOSE framework central theme is focused on the nature and connections of the structural elements that make up an explanation. While NOSE has a formative function (i.e., it is used to compare explanations to canonical science), it also emphasizes the relevance of these elements (pieces of knowledge, inferences, observations, lawlike statements, causal links, etc.) to the phenomenon-to-be-explained. Furthermore, NOSE examines how these elements are connected together in the construction of a scientific explanation. Here, it should be noted that NOSE dismisses the notion of absolute truth. Thus, a scientific explanation is not about "showing" which of a set of pieces of knowledge, inferences, laws, conditions is "true," but rather connecting *all* and *only* the relevant elements in a contextual, meaningful, and coherent/logical way. Thus, NOSE views the structural elements in explanations as content- and context-dependent (i.e., dependent on the phenomenon-to-be-explained, as well as, on the context in which explanations are generated).

Third, NOSE assesses explanations based on a continuous rather a dichotomous spectrum. Scientific explanations can be adequate, mostly adequate, partially adequate, or inadequate. In doing so, NOSE aims to shift assessment of explanations from the current traditional Correct/Incorrect ways of assessment. Additionally, the quality of a scientific explanation is dependent on the scientific phenomenon being explained. Fourth, NOSE framework follows a pragmatic approach to studying scientific explanations. Similar to Weber et al.'s (2013) approach, NOSE views philosophical models of scientific explanations as tools to assessing scientific explanations (i.e., different combinations of types of explanation).

Participant perceptions and NOSE framework. An examination of the criteria used by participants to assess the "goodness" of explanation reveals an interesting correspondence with the NOSE framework in several ways. First, as Figure 4.29 shows, the major common criteria that were emphasized both within and across the three participant groups correspond, to a significant extent, to pragmatic (context-dependent) nature emphasized by NOSE. More teachers and scientists than students emphasized the importance of constructing a learner-appropriate explanation as a criterion of a good scientific explanation. This criterion might stem from the fact that these participants function within the context of teaching and learning, which posits a concern for the audience (e.g., a concern for the needs of K-12 students in the case of teachers or undergraduate students in the case of scientists). When considering the practices that science teachers use to promote scientific explanation with students, helping students make sense of the explanations is a major focus.

The results further show that for a majority of participants in all three groups an explanation comprises a clear observation/description (the What-part) that is supported by scientific information (the Why-part) in the form of pieces of knowledge, scientific laws, scientific inferences, etc.. The explanation should present logical and coherent connections to show how laws and other related scientific information further support observations used. In a sense, participants agree with the NOSE framework that explanations should, at least, have pieces of knowledge, inferences, and laws. Some, though not the majority of participants, added a causal connection to the aforementioned elements. The fourth common criteria emphasized by a majority of participants in all three groups was that an explanation should be correct or accurate. This criterion is also aligned with the formative function of the NOSE framework, though it does not imply absolute truth.

It should be noted that the use of scientific information relevant to the phenomenon (and dismissing irrelevant information) was a major criterion of a good scientific explanation mentioned by all scientists, but a minor criterion for teachers and students. When judging peer and other participant group explanations, participants emphasized not only the importance of the existence of content knowledge in a good explanation, but more importantly, the nature (i.e., relevance and accuracy) of these facts to the phenomenon-to-be-explained. Finally, participants believed that explaining *all* related ideas (i.e., depth of explanation) to be an important criterion of a good scientific explanation. This overlaps with the dimension of completeness emphasized by the NOSE framework where an adequate explanation accounts for *all* relevant components of the phenomenon-to-be-explained. Often times, explanations were judged by other participants as incomplete for merely not explaining further some ideas mentioned. For example, not explaining why bubbles adhere to the raisins in the Dancing Raisins phenomenon was a common example mentioned by participants while assessing other participants' explanations.

For the most part, the major criteria for assessing explanations that were invoked by a majority of participants within and across the three groups overlapped with the NOSE framework and its structural elements. On the other hand, it could be argued that some criteria deployed by smaller groups of participants when assessing explanations do not overlap with the NOSE framework. For example, in Figure 4.30 some participants believed that adding examples related to everyday life strengthens an explanation. Additionally, some students and teachers believed that "mathematical equations" support observations used in a given explanation. Further criteria such as "clarity", "comprehensibility", "bias" and "depth of explanation" were considered by participants but are not directly emphasized in the NOSE framework. While NOSE framework does emphasize the importance of a clear and comprehensible explanation, other criteria (such as the use of examples, the depth of an explanation) derive from the social context within which an explanation is developed, while other minor elements are context and learner-dependent – two important elements emphasized in the NOSE framework. Finally, some participants from at least two groups considered objectivity as a criterion for a good scientific explanation. This could be considered as using clear, accurate, and relevant language, and identifying and addressing bias (to the extent possible).

CHAPTER 5: DISCUSSION AND IMPLICATIONS

This study aimed to, first, propose a domain-specific—namely in the physical sciences, framework that is designed for assessing scientific explanation in the science classrooms: The Nature of Scientific Explanation (NOSE) framework. Second, the study aimed to understand how students, science teachers, and scientists perceive the nature, and assess the quality of explanations in relation to NOSE. The study, thus, served to validate the usefulness of NOSE framework in the context of science teaching and learning. The researcher closely examined the explanations of two scientific phenomena that were generated by participant students, teachers and scientists. In addition, participants' perceptions of scientific explanation along with the criteria they drew upon when judging the "goodness" of explanations were examined. Participant perceptions and the criteria they used were compared with those characteristic of NOSE framework. Even though inferences were limited to the 30 participants in this study, the findings are promising and have important implications.

It is important to mention that because of the self-selected nature of participants and the relatively small sample size, this study cannot lay claims to generalizability. The participants were not necessarily representative of a larger group of freshmen college students, science teacher, and practicing scientists. To be sure, the study was exploratory in nature and did not aim for generalizability. Nonetheless, the results obtained were valuable in shedding light on the appropriateness of the expectations for using NOSE framework to examine scientific explanations.

Discussion of the Findings

Structure of Explanation. A NOSE framework analysis showed that participant scientists did significantly 'better' than teachers, who in turn did better than students. From the perspective of participants, 63% of scientists' explanations were assessed as complete and adequate (or good) explanations, 40% of teachers presented complete and adequate explanations, and only 13% of students did. What is more, scientists had more adequate scientific explanations, from a NOSE perspective, in the sense of providing more relevant and accurate structural elements. A NOSE framework analysis showed that the participants' explanation maps demonstrated similarities and differences across the three groups. Mainly, all participants used observations, inferences and pieces of knowledge along with laws and lawlike statements to explain why a phenomenon occurred. Yet, scientists' explanations included more pieces of knowledge and lawlike statements, which were relevant and accurate and/or based on prior content knowledge compared to students' and teachers' explanations. On the other hand, students and teachers made significantly more observations and inferences than scientists, and used these inferences to support their observations.

An important aspect of this study is the unique role that each participant group played. More specifically, while it was found, by participants, that scientists' explanations were the 'best', teacher participation was integral in highlighting contextual factors associated with explanation. In addition, student participation was important for whether or not the NOSE framework analysis placed realistic expectations on K-12 construction and assessment of scientific explanations.

Another finding was that, compared to scientists, students and teachers relied more on simple cause-effect relationships. This is consistent with prior studies on causality in explanation in science education (e.g., Braaten & Windschitl, 2010; Grotzer, 2003; Perkins & Grotzer, 2005). These studies also found that science teachers often faced challenges with the construction and assessment of causal explanations in the science classroom. In particular, in alignment with the present findings, Grotzer and Perkins found that teachers and students are only able to produce simple, linear cause-effect relationships to explain phenomena. In this study, it was evident that teachers and students had more difficulty generating more complex causal mechanistic explanations than scientists, while the latter heavily relied on causal mechanistic processes and interactions in their explanations. What is more, past research has emphasized the importance of adequate higher-order causal relationships. For instance, Grotzer and Basca (2003) argued that when students and teachers were given the proper training to develop appropriate causal links, their understanding of scientific phenomena was improved. Nonetheless, studies in science education often say little about the quality of the elements that participants use to make explanations and how scientists tend to construct explanations. This was another finding in the present study: CM explanations generated by participants were, by and large, adequate explanations in which the explainers used multiple cause-effect relationships in the form of causal process and causal connections. Also, scientists produced the majority of these CM (including CMDN and CMIS) explanations.

The quality of explanation elements is not addressed by almost all prior studies, mainly because these studies were limited by their analytical frameworks. Past research on the teaching and learning, as well as assessment, of learners' scientific explanation has often resorted to models that were, at best, peripherally relevant to the topic, such as Toulmin's (or a modified version of Toulmin's) model of argumentation, without necessarily making a convincing case that arguments are some type of explanations. This is a critical aspect. An appropriate examination of the nature of elements that make up an explanation, first, enables science education researchers to gain a better understanding of the nature of students' scientific explanations; and, second, provides a foundational approach to assess whether student-constructed ‘explanations’ can be considered explanatory.

Additionally, the present results highlight students’ perceptions of nature of science (i.e., there is a right/wrong way of doing science) and scientific knowledge (i.e., scientific knowledge is objective or “absolutely true”). As will be discussed later in this chapter, participants’ views of nature of science could be related to their perceptions of what explanations are. In addition, these findings could explain some of the observed shortcomings in students’ explanations. For example, while participants from all three groups emphasized the importance of having pieces of knowledge (often referred to by participants as content knowledge) and laws in an explanation, students were not as explicit about the need for making coherent and logical connections between pieces of knowledge and laws as were scientists and teachers. What is more, students tended to focus more on observations and the connection between observations and other structural elements. Additionally, while all participants acknowledged the role of explanation to provide understanding, students emphasized objectivity and “truth.” Teachers, on the other hand, emphasized an “appropriate level of correctness,” which highlights the pedagogical lens with which science teachers view, and in turn, assess explanations. Teachers also considered that explanations are contextual and learner dependent – a criterion aligned with the pragmatic notion of explanation emphasized in NOSE framework. Finally, scientists seemed to recognize the fact that scientific information alone (prior knowledge, scientific data, etc.) is not sufficient to provide understanding of a phenomenon. Instead, scientists emphasized the importance of relevance rather than mere correctness of prior knowledge, and the importance of logical coherence and flow among relevant elements of an explanation.

The literature on explanation (e.g., Brewer et al., 2000; Colombo, 2017; Erduran, Mestad & Kolstø, 2016; Zangori & Forbes, 2015; Simon & Osborne, 2004; Tang, 2016) has often argued that explanations created by scientists should serve as a reference or a standard for teacher and student explanations. A detailed examination of how students, teachers, and scientists' explanations actually differ, nonetheless, is rarely articulated in the literature. Furthermore, philosophical models of explanation were originally developed through an examination of explanations produced by scientists. The present study shows that scientists, teachers and students share a lot of similarities in how they construct their explanations. However, they differ in some key dimensions. To start with, as would be expected, scientists had more elaborate prior content knowledge, which allowed them to explore aspects of the scientific phenomenon that were not possible in the case of students and teachers. This finding is supported by past studies on novices and experts (e.g., Hmelo-Silver & Pfeffer, 2004; Grosslight, Unger, Jay, & Smith, 1991). For example, Grosslight et al. compared expert and novice conceptions of models and their use in science, and examined the criteria that middle and high school students used to decide whether or not specific items were scientific models. They further interviewed experts for comparison and found that students' conceptions of models were consistent with a naïve realist epistemology, in which students thought that models are physical copies of reality. On the other hand, they found that experts' ideas were consistent with a constructivist framework, where they viewed models as somewhere between abstraction and reality. In the present study, results in this regard were similar to the case of generating models.

Another equally interesting finding is related to teleology and anthropomorphism. It is noteworthy that teleology may be considered a special cause of anthropomorphism (as suggested by Hempel, 1965). To that end, there is a consensus among researchers that the biological sciences can evoke teleology and anthropomorphism; whereas, physical scientists in particular, explicitly reject them and caution against using them when explaining natural phenomena (e.g., Kelemen, Rottman, & Seston, 2013; Tamir & Zohar, 1991). Stemming from this school of thought, and since for now NOSE framework focuses on explanations in the physical sciences, teleological and anthropomorphic statements were not considered a separate kind of explanation as DN, IS, or CM explanations were. Instead, teleologic and anthropomorphic statements, when present, were considered as structural elements of the explanation at hand. In the present study, the use of teleology and/or anthropomorphism revealed interesting findings in relation to the structure of a scientific explanation.

While constructing their DR and CIJ explanations, teachers tended to use teleological and/or anthropomorphic statements strikingly differently from students and scientists. In fact, teachers seemed to use these statements as a pedagogical tool to help with understanding the phenomenon at hand. This finding is aligned with Hempel's (1965) philosophical work, which suggests that teleological explanations make learners feel that they understand the phenomenon-to-be-explained because these explanations are provided in terms of purposes and intentions. The latter better fit the way people are used to view their own behavior. What is more, in the majority of teachers' use of anthropomorphic statements, there did not seem to be any implication of anthropomorphic reasoning. For example, when Tarra, a participant teacher, used the statement: "the molecules of water *want* [emphasis added] to be together," she did not imply that these molecules actually had a wanting or a desire to *be together*. Much of the literature on teleology and anthropomorphism in science education focuses on students' rather than teachers' views (e.g., Bartov, 1978; Crannell, 1954; Kallery, 2001; Tamir & Zohar; 1991). However, in one study, Kallery and Psillos (2004) examined teachers' views on the use of animism and anthropomorphism. They found that teachers used animism and anthropomorphism both knowingly and unknowingly, and that they justified their conscious use of these kinds of statements as appropriate to the learners' level and content knowledge.

No doubt, some of the present findings corroborate prior research findings, especially, in how each of the participant groups think when they explain. Nonetheless, prior studies in science education fall short in considering aspects specific to explanations. For instance, prior studies do not consider the context of teaching and learning when constructing explanations. Rather, it is assumed that the "best" form or explanation within a science context is that which replicates scientists' explanations. Moreover, prior studies in science education do not examine the different types of explanations for different natural phenomena.

Perceptions of, and criteria to assess, explanation. Like any other construct, explanation can have different meanings for different people in different contexts. In the process of doing research on explanation, researchers typically narrow the meaning of explanation for practical purposes. Yet, difficulties arise when researchers in science education generalize their definition of explanation across participants and contexts (e.g., Kesonen, et al., 2017; Yao, et al., 2016). One important aspect that should be taken into account when examining scientific explanation is the context of science teaching and learning related to students' and teachers' notions of explanation. The context factor allows for determining what is relevant to an explanation and what is not. For example, results from the DR explanations showed that scientists tended to explain more aspects of the DR phenomenon than did students and teachers. In particular, scientists considered explanations of why the bubbles adhered to the raisins, or why the bubbles popped at the top as relevant to explaining the behavior of the raisins. On the other hand, students and teachers generally seemed to be mainly concerned with explaining why the raisins behaved the way they did. In fact, some teachers even considered the aforementioned factors (i.e., the bubbles adhering to the raisins, or the popping of bubbles at the top) to be irrelevant to the DR explanation.

Another factor that the present study highlights is that the type(s) of explanations produced mainly depends on the nature of the phenomenon-to-be-explained. For example, results in this study showed that causality was significantly more evident in participants' DR explanations than in their CIJ explanations. In fact, 17 of the 30 CIJ explanations were either Causal, CDN, or CMDN explanations as compared to 28 DR explanations. This could be explained by the different natures of each phenomenon: it was more adequate to use cause-effect relationships to explain how the bubbles adhering to the raisins *caused* the raisins to float. However, the rising of the water in the CIJ scenario could be mainly explained by subsuming the phenomenon under laws and lawlike statements, and still provide an adequate explanation. Furthermore, even though Sam, a participating scientist, tended to use laws and lawlike statements of probabilistic nature to explain the DR phenomenon, his explanation was not adequate, mainly because he used probabilistic laws to explain a fairly deterministic phenomenon.

This study also showed that participants indeed have perceptions about what explanations are. In particular, students tend to think of explanation as a "true" answer to a *why*-question based on observations. However, teachers and scientists tended to perceive explanation as a testable and verifiable tool that provides understanding. As noted earlier, this can possibly be related to the views participants hold about the nature of science and scientific knowledge. Such views need to be considered prior to engaging students, teachers or scientists in explanation. The present study shows, interestingly, that teachers were the only participant group to show somewhat parallel views of explanation to those of Toulmin's views on argumentation. Nonetheless, unlike Toulmin, teachers as well as students and scientists, could not eliminate context-dependent and learner-dependent aspects of explanation, as was noted in Chapter 4. For the case of teachers, such a finding speaks to the current training that they receive where Toulmin's (Claim-Evidence-Reasoning) model of argumentation is predominantly applied across various subjects and different constructs in the school setting.

Finally, participants seemed to acknowledge some fundamental demarcation criteria between what is explanatory and what is not. This was especially evident in teachers' and scientists' explicit distinction between *what*- and *why*-questions. Hence, the present study provided compelling evidence against prior studies in which researchers, in many cases, have considered *all* students' answers as explanations (e.g., Braaten & Windschitl, 2010; Forbes et al., 2014; Kesonen et al., 2017; Mestad & Kolstø, 2016; Southard et al., 2017).

More important were the criteria that participants used to assess explanations. NOSE framework emphasizes context-dependent aspects of explanation instead of a sole focus on the structural elements of an explanation. This turned out to be an important aspect of the findings of this study. When participants assessed the adequacy and completeness of their own explanations, of other participants from their group, and of participants from the other two groups, a key criterion observed was related to the context in which an explanation was produced. In particular, participants from all groups emphasized that a "good" explanation is one that is appropriate to the learner—the target of the explanation.

What is more interesting, and what is probably the most powerful finding in establishing the usefulness and validity of the NOSE framework in the present study, was the fact that starting with whatever criteria they had for assessing the "goodness" of explanation, most participants across all three groups judged as “best” or “complete” or “good” the explanations made by participant scientists. Hence, the present study validated that the reference standard for explanation were ones made by scientists, followed by ones made by teachers, and very few made by students. This finding aligns with the reforms’ emphasis on the need for attaining instructional outcomes for science students within authentic scientific practice (e.g. NRC, 2000). Additionally, judging 40% of teachers’ explanations as “best” could also be explained by the fact that teachers have pedagogical expertise that make them communicators of complicated scientific ideas to learners.

In conclusion, the present study highlighted the need articulated by many researchers in science education to understand additional aspects specific to scientific explanation. The study highlighted the importance of not only the structural elements that make up a scientific explanation, but also the connectedness of these elements within the context of teaching and learning. There is no doubt that social, and even cultural, aspects come into play when constructing explanations be it for scientists, teachers or students. These aspects are the foundation of criteria generated by participants when assessing explanations. Identifying such criteria is pertinent to understanding that what counts as an adequate scientific explanation changes by students’ level, prior knowledge and other factors determined by the general context.

Perceptions and the use of NOSE. In essence, NOSE framework, in its current emergent stage, provides an adaptive schema of explanations that is grounded in philosophical models and approaches of scientific explanation. NOSE emphasizes the idea that in some science topics, events can be explained by referring to general laws (the DN model), highly probable laws (the IS model), and/or causal mechanistic processes (the CM model) within a pragmatic approach that considers students’ levels and their prior knowledge, as well as the context of learning. Without such a framework, problems seen in prior studies of using other frameworks, or no frameworks at all, lie in the overgeneralization of what actually counts as explanatory without due consideration to the many contextual variables that exist. An interesting aspect of this study was that participants' views of the "best" explanation corroborated the analysis from a NOSE framework perspective.

Implications for Practice

Although explanation in science education has been emphasized by researchers, scholars, and teachers as a goal for science and a tool for attaining understanding, the complexities of explanation have been overlooked. The researcher, thus, argues that through the development of the NOSE framework in the present study, attention is now directed more on the nature of explanation. This includes (a) gaining an understanding, and utilizing the different types, of explanations from a philosophical and theoretical perspective; (b) examining various structural elements that make up different types of explanations and how these elements interconnect; and (c) clarifying contextual and learner variables that specify relevance, completeness, and depth of explanations.

In teaching the practice of explanation within an inquiry-based context, teachers need to not only focus on the general structures of an accepted construct, be it argument, explanation or something else, but they need to focus on the quality of the structures produced within the relevant context they are being produced. In other words, teachers need to address what *kind* of observations, pieces of knowledge, inferences, laws, necessary conditions, etc. would be considered relevant and accurate within a given context; in addition to the ways by which these structural elements are interconnected during the process of developing explanations.

Implications for Research

This study has shown that there are various issues overlooked by researchers in science education regarding explanation. The most predominant of these is the absence of a clear set of guidelines and modalities that are unique to scientific explanation. The study shows that there are criteria specific to explanation that are detrimental to its structural validity. Hence, using frameworks that are not originally developed to examine explanations tends to overlook the underlying criteria specific to explanations. In other words, science educators need to acknowledge the criteria specific to explanation when examining learner-constructed explanations. NOSE framework is among the first attempts in science education that aimed to develop a functional framework of scientific explanation guided by the underlying philosophical models that is useful for K-12 science teaching and learning. As discussed in Chapter 1, NOSE framework in the present study is intended mainly for use by science education researchers. It is sought to provide researchers with a tool to enable them to meaningfully assess students' constructed explanations in different settings. Additionally, NOSE framework proposed here is not set in stone, but rather emergent. As discussed in Chapter 4, findings from the present study revealed that NOSE framework overlooked some elements that were observed in participants' explanations (e.g., examples, mathematical equations) and some criteria that were considered by participants (such as simplicity, connectedness to real life, comprehensibility). Through continued use and analysis of the NOSE framework, empirical data might suggest the need for additional elements, categories, and perhaps, types of explanations.

Implications for Future Research

During the process of this study, the researcher recognized that there needs to be a reorientation of *how* research is done on the practice of explanation in the science classroom. In addition, philosophical models of explanation constitute robust support in the construction of a framework unique to explanation, and need to be the basis of a meaningful examination of explanation. There is no doubt that findings in this study emphasize the need to understand further the types of explanations constructed by explainers, and the underlying assumptions they hold about explanation. It will also be useful to investigate whether or not such assumptions differ significantly from one classroom context to another. To that end, further understanding of the contexts in which scientists construct their explanations is also needed to create a better understanding of the nature of scientific explanation.

Hence, four major directions emerge from this work as far as future research studies go. The first direction seeks to explore ways by which NOSE framework can be adapted into a pedagogical framework that teachers can use to facilitate adequate construction, and meaningful assessment of scientific explanations within the various contexts of teaching science. Such a framework would be based on the philosophically-grounded NOSE framework and the criteria developed from participants generated from this study. Questions that arise from this line of research include: Does the use of such a framework improve teachers' understanding of explanation and, in turn, improve students' abilities to construct adequate explanations? What difficulties do teachers face when using such a framework? How does the use of this framework in pre-service classrooms affect their perceptions and assessment of scientific explanation? Guided by the NOSE framework, an immediate study that can be conducted in this direction involves engaging pre-service elementary teachers, in a science education methods classroom, in reflecting on their own explanations and developing explanations that are age-appropriate.

The second direction emphasizes a better understanding of the nature of explanation within the field of science. It seeks to understand how different disciplinary cultures in science influence the construction of scientific explanations. Some related questions include: Do scientists explain differently from one subject domain to another? If so, how do these explanations compare within different domains? Do the different K-12 subject domains (i.e., physics, chemistry, biology, etc.) reflect the processes by which scientists in these respective domains construct explanations? If not, how might this approach improve science teaching in the classroom?

The third direction seeks to examine the factors that affect the quality of explanations constructed by students in science classrooms. One way to approach this is to look at the impact of students' prior knowledge on their explanation construction across the elementary, middle and high school levels. Related questions include: In what ways do students' scientific explanations change as the students move through the curriculum? How do factors such as, socio-economic status, language fluency, and cultural background impact students' construction of scientific explanation?

The fourth direction seeks to understand whether the practice of explanation construction within classroom contexts enhances students' understanding of the nature of science. If so, then what perceptions do students hold about the nature of science after engaging in meaningful explanation construction? Do these students' perceptions differ from those of students who do not participate in such practice? In addition, when do students find explanations adequate (or satisfying) and does this influence their perspective on how they identify and assess an explanation as well as their understanding of the nature of science?

The above directions are focused on developing a well-rounded understanding of explanation in science education. Implications from such future research are not only limited to the explanation field but address multiple issues in teacher education and student learning. It does not escape the researcher's attention that a discourse analysis can also apply to this work brining more insight to the conceptualization of scientific explanation. These implications also aim to further develop a deeper understanding of the nature of scientific explanation within the context of science education research. Thus, a thorough and reflective exploration of the above issues (along with other issues that will emerge) can help inform research on scientific explanation in science education.

REFERENCES

Abi-El-Mona, I., & Abd-El-Khalick, F. (2011). Perceptions of the nature and 'goodness' of argument among college students, science teachers, and scientists. *International Journal of Science Education*, *33*, 573-605.

ACARA. (2015*). Australian Curriculum, Assessment and Reporting Authority*. Retrieved from: https://www.australiancurriculum.edu.au/f-10-curriculum/science/.

Achinstein, P. (1984). The pragmatic character of explanation. In Asquith, P., & Kitcher, P. (Eds.), *Philosophy of science*, East Lansing, MI. pp. 275–292.

Alameh, S., & Abd-El-Khalick, F. (2018). Towards a Philosophically Guided Schema for Studying Scientific Explanation in Science Education. *Science & Education*, *27*, 831-861.

Bartov, H. (1978). Can students be taught to distinguish between teleological and causal explanations, *Journal of Research in Science Teaching, 15,* 567-572.

Beebee, H., Hitchcock, C., & Menzies, P. (Eds.). (2009). *The Oxford handbook of causation*. Oxford University Press.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education, 22*, 797–817.

Berland, L. K. and Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26-55.

Beyer, C.J., and Davis, E.A. (2008). Fostering second graders' scientific explanations: A beginning elementary teacher's knowledge, beliefs, and practice. *The Journal of the Learning Sciences*, *17*(3), 381-414.

Braaten, M., & Windschitl, M. (2010). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education, 95*, 639–669.

Brewer, W. F., Chinn, C. A., & Samarapungavan, A. (2000) Explanation in scientists and children. In F. C. Keil and R.A. Wilson (eds*.) Explanation and cognition.* Cambridge, MA. MIT Press.

Brigandt, I. (2016). Why the Difference Between Explanation and Argument Matters to Science Education. *Science and Education*, *25*(3-4), 251-275.

Bromberger, S. (1966). Why-questions. In R. G. Colodny (Ed.), *Mind and cosmos: Essays in contemporary science and philosophy* (pp. 86–110). Pittsburg: University of Pittsburg Press.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

Chi, M. T. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *The journal of the learning sciences*, *14*, 161-199.

Colombo, M. (2017). Experimental philosophy of explanation rising: The case for a plurality of concepts of explanation. *Cognitive Science*, *41*, 503-517.

Crannell, C. N. (1954). Responses of college students to a questionnaire on animistic thinking. *Scientific Monthly, 78*, 54–56.

Craver, C. F. (2007). Explanation and Causal Relevance. In C. F. Craver (Ed.), *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. (pp. 21-62). Oxford: Clarendon Press.

Dagher, Z., & Cossman, G. (1992). Verbal explanations given by science teachers: Their nature and implications. *Journal of Research in Science Teaching*, *29*(4), 361-374.

De Andrade, V., Freire, S., and Baptista, M. (2017). Constructing Scientific Explanations: A System of Analysis for Students' Explanations. *Research in Science Education*, 1-21.

de Carvalho, A. M. P., & Paulo, S. (2004). Building up explanations in physics teaching. *International Journal of Science Education*, *26*(2), 225-237.

De Regt, H. W., Leonelli, S., & Eigner, K. (Eds.). (2009). *Scientific understanding: philosophical perspectives*. University of Pittsburgh Press.

De Vries, E., Lund, K., & Baker, M. (2002). Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *The journal of the learning sciences*, *11*(1), 63-103.

Delen, I., & Krajcik, J. (2018). Synergy and students' explanations: Exploring the role of generic and content-specific scaffolds. *International Journal of Science and Mathematics Education*, *16*(1), 1-21.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2011). *The Sage handbook of qualitative research*. Thousand Oaks, CA: Sage.

diSessa, A. (1986). Knowledge in pieces. In G. Forman & P. Pufal (Eds.). *Constructivism in the computer age.* Hillsdale, NJ: Erlbaum.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science education*, *84*(3), 287-312.

Eberbach, C., & Crowley, K. (2009). From everyday to scientific: How children learn to observe the biologist's world. *Review of Educational Research, 79*(1), 39–68.

Erduran, S., Simon, S., and Osborne, J. (2004). Tapping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, *88*, 915-933.

Feigl H. The 'orthodox' view of theories: remarks in defense as well as critique. In: Radner M, Winokur S, editors. Minnesota studies in the philosophy of science. Minneapolis: University of Minnesota Press; 1970. pp. 3–16.

Forbes, C., Lange, K., Möller, K., Biggers, M., Laux, M., & Zangori, L. (2014). Explanation-Construction in Fourth-Grade Classrooms in Germany and the USA: A cross-national comparative video study. *International Journal of Science Education, 36*, 2367- 2390.

Ford, D. (2005). The challenges of observing geologically: Third graders' descriptions of rock and mineral properties. *Science Education*, *89*, 276–295.

Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, *71*(1), 5-19.

Gardiner, P. (1959). *The Nature of Historical Explanation*, Oxford: Oxford University Press.

Gilbert, J. K., Boulter, C., & Rutherford, M. (1998). Models in explanations, Part 1: Horses for courses?. *International Journal of Science Education*, *20*(1), 83-97.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, *69*(3), 342-353.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science teaching*, *28*(9), 799-822.

Grotzer, T. (2003). Learning to understand the forms of causality implicit in scientiWcally accepted explanations. *Studies in Science Education, 39*, 1–74.

Grotzer, T. A., & Basca, B. B. (2003). How does grasping the underlying causal structures of ecosystems impact students' understanding?. *Journal of Biological Education*, *38*(1), 16-29.

Haefner, L.A., and Zembal-Saul, C. (2004). Learning by doing? Prospective elementary teachers' developing understandings of scientific inquiry and science teaching and learning. *International Journal of Science Education*, *26*, 1653-1674.

Halls, J. G., Ainsworth, S. E., & Oliver, M. C. (2018). Young children's impressionable use of teleology: the influence of question wording and questioned topic on teleological explanations for natural phenomena. *International Journal of Science Education*, *40*(7), 808-826.

Hausman, D. M. (1998). *Causal asymmetries*. Cambridge: Cambridge University Press.

Hempel, C. (1965). Aspects of scientific explanation. In C. Hempel (Ed.), *Aspects of scientific explanation, and other essays in the philosophy of science.* (pp. 331-489). New York, NY: Free Press.

Hempel, C. & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*, 135 – 175.

Hempel, C. G. (1962). Explanation in Science and in History. *Frontiers of science and philosophy*, 7-33.

Hempel, C. G. (2001). *The philosophy of Carl G. Hempel: studies in science, explanation, and rationality*. Oxford University Press.

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive science*, *28*(1), 127-138.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *38*, 663-687.

Horn, R. E. (1998). *Visual language: Global communication for the 21st century.* Bainbridge Island, WA: MacrovVU, Inc.

Hume Studies, Volume 11, Number 1, April 1985, pp. 94-108 (Article)

Humphreys, P. W. (1989). Scientific explanation: The causes, some of the causes, and nothing but the causes. In: Kitcher P, Salmon W (eds) Scientific explanation. University of Minnesota Press, Minneapolis, pp 282–306

Ivany, J. G., & Oguntonade, C. B. (1972). Verbal explanation in physics classes. *Journal of Research in Science Teaching*, *9*(4), 353-359.

Jang, J. Y., & Hand, B. (2017). Examining the value of a scaffolded critique framework to promote argumentative and explanatory writings within an argument-based inquiry approach. *Research in Science Education*, *47*(6), 1213-1231.

Jungwirth, E. (1979). Do Students Accept Anthropomorphic and Teleological Formulations as Scientific Explanations?. *Journal of College Science Teaching*, *8*(3), 152-55.

Kallery, M. (2001). Early-years educators' attitudes to science and pseudo-science: The case of astronomy and astrology. *European Journal of Teacher Education, 24*(3), 329–342.

Kallery, M., & Psillos, D. (2004). Anthropomorphism and animism in early years science: Why teachers use them, how they conceptualise them and what are their views on their use. *Research in Science Education*, *34*(3), 291-311.

Kampourakis, K., & Zogza, V. (2008). Students' intuitive explanations of the causes of homologies and adaptations. *Science & Education*, *17*(1), 27-47.

Kampourakis, K., Pavlidi, V., Papadopoulou, M., & Palaiokrassa, E. (2012). Children's teleological intuitions: What kind of explanations do 7–8 year olds give for the features of organisms, artifacts and natural objects? *Research in Science Education*, *42*, 651-671.

Kampourakis, K., Silveira, P., & Strasser, B. J. (2016). How do preservice biology teachers explain the origin of biological traits?: A philosophical analysis. *Science Education, 100*, 1124-1149.
Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, *142*(4), 1074.
Kesonen, M. H. P., Asikainen, M. A., & Hirvonen, P. E. (2017). Light Source Matters–Students' Explanations about the Behavior of Light When Different Light Sources are used in Task Assignments of Optics. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(6), 2777-2803.
Kim, D., & Benbasat, I. (2006). The effects of trust-assuring arguments on consumer trust in Internet stores: Application of Toulmin's model of argumentation. *Information Systems Research*, *17*, 286-300.
Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation*. (pp. 410-499). Minneapolis, MN: University of Minnesota Press.
Kokkonen, T., & Mäntylä, T. (2018). Changes in university students' explanation models of dc circuits. *Research in Science Education*, *48*(4), 753-775.
Kokkonen, T., & Mäntylä, T. (2018). Changes in university students' explanation models of dc circuits. *Research in Science Education*, *48*(4), 753-775.
Laplace, P. S. (1951). *A philosophical essay on probabilities*. Translated by Truscott, F. W., & Emory, F. L., New York: Dover Publications.
Lawson, A. E., Drake, N., Johnson, J., Kwon, Y. J., & Scarpone, C. (2000). How good are students at testing alternative explanations of unseen entities? *The American Biology Teacher*, 249-255.
Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of research in science teaching*, *39*(6), 497-521.
Lipton, P. (2004). What good is an explanation?. In J. Cornwell (Ed.), *Explanations: styles of explanation in science (pp. 1-21)*. Oxford: Oxford University Press.
LOMCE. (2015*). Spanish Law for the Improvement of Quality of Education*. Ministerio de Educación de España.
Mayes, G. R. (2010). Argument explanation complementarity and the structure of informal reasoning. *Informal Logic*, *30*(1), 92-111.
McNeill, K. L., and Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, *45*(1), 53-78.
McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*, *15*, 153-191.
Mestad, I., & Kolstø, S. D. (2017). Characterizing Students' Attempts to Explain Observations from Practical Work: Intermediate Phases of Understanding. *Research in Science Education*, *47*(5), 943-964.
Metz, K. (1991). Development of explanation: Incremental and fundamental change in children's physics knowledge. *Journal of Research in Science Teaching, 28*, 785-797.

Meyer, K., and Woodruff, E. (1997). Consensually driven explanation in science teaching. *Science Education*, *80*, 173-192.
Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*, New York: Harcourt, Brace and World.
National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13165.
NCE. (2015). *National Curriculum of England.* United Kingdom, Department of Education. Retrieved from: https://www.gov.uk/government/publications/national-curriculum-in-england- secondary-curriculum.
NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Retrieved from http://www.nextgenscience.org/.
Norris, S. P., Guilbert, S. M., Smith, M. L., Hakimelahi, S., & Phillips, L. M. (2005). A theoretical framework for narrative explanation in science. *Science Education*, *89*(4), 535-563.
Osborne, J., and Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, *95*(4), 627-638.
Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of research in science teaching, 41*, 994-1020.
Papadouris, N., Vokos, S., & Constantinou, C. P. (2018). The pursuit of a “better” explanation as an organizing framework for science teaching and learning. *Science Education*, *102*(2), 219-237.
Parnafes, O. (2012). Developing explanations and developing understanding: Students explain the phases of the moon Using visual representations. *Cognition and Instruction, 30*, 359-403.
Peker, D., & Wallace, C. S. (2011). Characterizing high school students’ written explanations in biology laboratories. *Research in Science Education*, *41*, 169-191.
Perkins, D.N. & Grotzer, T.A. (2008). Dimensions of causal understanding: The role of complex causal models in students' understanding of science. Studies in Science Education, 41 (1), 117–165
Pisa, O. E. C. D. (2015). Draft science framework. *2014-07-17]. http://www. oecd.org/pisa/pisaproducts/Draft PISA 2015 Science Framework. pdf.*
Popper, K., 1959, *The Logic of Scientific Discovery*, London: Hutchinson.
Railton, P. (1978), A Deductive-Nomological model of probabilistic explanation. *Philosophy of Science, 45*, 206–226.
Ruiz-Primo, M.A., Li, M., Tsai, S.P., and Schneider, J. (2010). Testing one premise of scientific inquiry in science classrooms: Examining students’ scientific explanations and student learning. *Journal of Research in Science Teaching*, *47*(5), 583-608.
Russ, R. S., Scherr, R. E., Hammer, D., and Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*, 499-525.
Ryder, J. (2001). Identifying science understanding for functional scientific literacy. Studies in Science Education, 36, 1 – 46
Sadler, T. (2006). Promoting discourse and argumentation in science teacher education. *Journal of Science Teacher Education*, *17*, 323-346.

Salmon, W. (1989). Four decades of scientific explanation. In P. Kitcher, P. & W.C. Salmon (Eds), *Scientific explanation*. (pp. 3-219). Minnesota, MN: University of Minnesota Press.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.

Salmon, W. C. (1998). *Causality and explanation*. New York: Oxford University Press.

Salmon, Wesley C. (1971), "Statistical Explanation", in Wesley C. Salmon et al., *Statistical Explanation and Statistical Relevance*. Pittsburgh: University of Pittsburgh Press, 29– 87.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, *12*(1), 5-51.

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and instruction*, *23*(1), 23-55.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, *88*, 345-372.

Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. Denzin, & Y. Lincoln (Eds), *Handbook of qualitative research* (2nd ed). (pp. 189-214). London: Sage Publications.

Scriven, Michael. 1959. Definitions, explanations, and theories. In *Minnesota Studies in the Philosophy of Science* Volume II, eds. H. Feigl, M. Scriven, and G. Maxwell. Minneapolis: University of Minnesota Press.

Sevian, H., & Gonsalves, L. (2008). Analysing how scientists explain their research: A rubric for measuring the effectiveness of scientific explanations. *International Journal of Science Education*, *30*(11), 1441-1467.

Simon, S., Erduran, S., and Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education*, *28*(2-3), 235-260.

Smith, B.O., & Meux, M.O. (1970). *A study of the logic of teaching.* Urbana, IL: University of Illinois Press.

Songer, N. B., Kelcey, B., & Gotwals, A. W. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *46*(6), 610-631.

Southard, K. M., Espindola, M. R., Zaepfel, S. D., & Bolger, M. S. (2017). Generative mechanistic explanation building in undergraduate molecular and cellular biology. *International Journal of Science Education*, *39*(13), 1795-1829.

Talanquer, V. (2018). Exploring mechanistic reasoning in chemistry. In Yeo, J., Teo, T.W., and Tang, K.S. (Eds.), *Science Education Research and Practice in Asia-Pacific and Beyond (*39-52). Singapore: Springer.

Talanquer,V. (2007). Explanations and teleology in chemistry education. *International Journal of Science Education*, *29*, 853-870.

Tamir, P., & Zohar, A. (1991). Anthropomorphism and teleology in reasoning about biological phenomena. *Science Education*, *75*(1), 57-67.

Tang, K.S. (2016). Constructing scientific explanations through premise–reasoning– outcome (PRO): an exploratory study to scaffold students in structuring written explanations. *International Journal of Science Education*, *38*(9), 1415-1440.

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

Treagust, D. F., & Harrison, A. G. (2000). In search of explanatory frameworks: An analysis of Richard Feynman's lecture'Atoms in motion'. *International Journal of Science Education*, *22*(11), 1157-1170.

Trommler, F., Gresch, H., & Hammann, M. (2018). Students’ reasons for preferring teleological explanations. *International Journal of Science Education*, *40*(2), 159-187.

Van Fraassen, Bas. (1980). *The Scientific Image.* Oxford: Clarendon Press.

Van Gelder, T. (2002). Argument mapping with reason! able. *The American Philosophical Association Newsletter on Philosophy and Computers*, *2*(1), 85-90.

Wang, C.-Y. (2014). Scaffolding Middle School Students’ Construction of Scientific Explanations: Comparing a cognitive versus a metacognitive evaluation approach. *International Journal of Science Education,* 37, 237-271.

Weber, E., Van Bouwel, J., De Vreese, L. (2013). *Scientific explanation*. NY: Springer.

Weinberg, R. A. (1985). The molecules of life. *Scientific American*, *253*(4), 48-57.

Woodward, J. (2003). *Making things happen: A theory of causal explanation.* Oxford: Oxford University Press.

Woody, A. I., (2013). How is the ideal gas law explanatory? *Science & Education*, *22*, 1563–1580.

Yang, H. T., & Wang, K. H. (2014). A teaching model for scaffolding 4th grade students’ scientific explanation writing. *Research in Science Education, 44*(4), 531-548.

Yao, J.X., Guo, Y. Y., and Neumann, K. (2016). Towards a hypothetical learning progression of scientific explanation. *Asia-Pacific Science Education*, *2*(1), 1-17.

Ye, L. R., P. E. Johnson. 1995. The impact of explanation facilities on user acceptance of expert systems advice. *Management Inform Systems Quart., 19*, 157–172.

Yeo, J., & Gilbert, J. K. (2014). Constructing a scientific explanation—A narrative account. *International Journal of Science Education*, *36*(11), 1902-1935.

Zangori, L., Forbes, C.T., and Biggers, M. (2013). Fostering student sense making in elementary science learning environments: Elementary teachers’ use of science curriculum materials to promote explanation construction. *Journal of Research in Science Teaching*, *50*, 989-1017.

Zangori, L., Forbes, C.T., and Schwarz, C.V. (2015). Exploring the Effect of Embedded Scaffolding Within Curricular Tasks on Third-Grade Students’ Model-Based Explanations about Hydrologic Cycling. *Science and Education*, *24*, 957-981.

Zuzovsky, R., & Tamir, P. (1999). Growth patterns in students' ability to supply scientific explanations: Findings from the Third International Mathematics and Science Study in Israel. *International Journal of Science Education*, *21*(10), 1101-1121.

APPENDIX A

SAMPLE FLYER AND SAMPLE INVITATION LETTERS

SAMPLE LETTER/EMAIL OF INVITATION
For the Participating Freshmen College Student

University of Illinois at Urbana-Champaign
Institutional Review Board # 19606

[Department of Curriculum and Instruction Letterhead]

(Date)

Dear Student,

My name is Sahar Alameh and I am a doctoral student in the Department of Curriculum and Instruction at the University of Illinois at Urbana Champaign. I am majoring in Science Education and I am currently starting to collect data for my dissertation research. I would like to invite you to participate in my research project, which aims at exploring students' explanations in science. This project will be supervised by Dr. David Brown, a professor in Science Education at the Department of Curriculum and Instruction.

I am interested in ways to help students develop the abilities and skills needed to construct and assess explanations in science. Such an instructional outcome is an indispensable goal of science. However, research in science education has shown that students face difficulties in constructing explanations in science. Therefore, this study is aimed at helping researchers find better ways to improve students' construction of scientific explanations.

Taking part in this project entails participating in two separate individual interviews about your ideas of everyday science phenomena. The interviews involve fun and exciting questions and are not aimed to evaluate your answers. There are no correct/wrong answers. We want to learn about the ways by which you explain phenomena and what you think about explanations.

Each interview is expected to last between 40 to 50 minutes. I will conduct the interviews, which will be videotaped with your permission. I will work with you to schedule the interviews at times that are most convenient to you.

If you participate in this study you will receive $20 cash value after completing the second interview.

Your participation in this project in completely voluntary, and you are free to withdraw at any time and for any reason without penalty. You are also free to refuse to answer any questions you do not wish to answer. Your decision to participate or not participate in this project will have no consequences and will not affect your relationship with the University of Illinois in any shape or form. Any information collected during the study will be kept in strict confidence. The results of the research will be presented in professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excerpts to us.

You will get to sign an informed consent form that details and ensures all your rights as a participant. If you wish, you can receive a copy or an executive summary of the research results after the project is completed.

If you have any questions about this research project, please contact Mrs. Sahar Alameh by telephone at 217-979-5471 (call or text) OR by email at alameh2@illinois.edu OR Dr. David Brown at debrown@illinois.edu. Their address location is the following:

Department of Curriculum and Instruction
University of Illinois at Urbana-Champaign
Education Building, 1310 S Sixth Street
Champaign, Illinois, 61820

I am hopeful you will agree to help with this research project.

Best Regard,
Sahar Alameh

SAMPLE LETTER/EMAIL OF INVITATION
For the Participating Secondary Science Teacher

University of Illinois at Urbana-Champaign
Institutional Review Board # 19606

[Department of Curriculum and Instruction Letterhead]

(Date)

Dear Student,

My name Is Sahar Alameh and I am a doctoral student in the Department of Curriculum and Instruction at the University of Illinois at Urbana Champaign. I am majoring in Science Education and I am currently starting to collect data for my dissertation research. I would like to invite you to participate in my research project, which aims at exploring students' explanations in science. This project will be supervised by Dr. David Brown, a professor in Science Education at the Department of Curriculum and Instruction.

I am interested in ways to help students develop the abilities and skills needed to construct and assess explanations in science. Such an instructional outcome is an indispensable goal of science. However, research in science education has shown that students face difficulties in constructing explanations in science. Therefore, this study is aimed at helping researchers find better ways to improve students' construction of scientific explanations.

Taking part in this project entails participating in two separate individual interviews about your ideas of everyday science phenomena. The interviews are not meant to evaluate your answers, but rather learn about the ways by which science teachers explain phenomena and what they think about explanations. This will greatly help us gain a better understanding for examining students' scientific explanations.

Each interview is expected to last between 40 to 50 minutes. I will conduct the interviews, which will be videotaped with your permission. I will work with you to schedule the interviews at times that are most convenient to you.

If you participate in this study you will receive $20 cash value after completing the second interview.

Your participation in this project in completely voluntary, and you are free to withdraw at any time and for any reason without penalty. You are also free to refuse to answer any questions you do not wish to answer. Your decision to participate or not participate in this project will have no consequences and will not affect your relationship with the University of Illinois in any shape or form. Any information collected during the study will be kept in strict confidence. The results of the research will be presented in professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excerpts to us.

You will get to sign an informed consent form that details and ensures all your rights as a participant. If you wish, you can receive a copy or an executive summary of the research results after the project is completed.

If you have any questions about this research project, please contact Mrs. Sahar Alameh by telephone at 217-979-5471 (call or text) OR by email at alameh2@illinois.edu OR Dr. David Brown at debrown@illinois.edu. Their address location is the following:

Department of Curriculum and Instruction
University of Illinois at Urbana-Champaign
Education Building, 1310 S Sixth Street
Champaign, Illinois, 61820

I am hopeful you will agree to help with this research project.

Best Regard,

Sahar Alameh

SAMPLE LETTER/EMAIL OF INVITATION
For the Participating Scientist

University of Illinois at Urbana-Champaign
Institutional Review Board # 19606

[Department of Curriculum and Instruction Letterhead]

(Date)

Dear Scientist,

My name is Sahar Alameh and I am a doctoral student in the Department of Curriculum and Instruction at the University of Illinois at Urbana Champaign. I am majoring in Science Education and I am currently starting to collect data for my dissertation research. I would like to invite you to participate in my research project, which aims at exploring students' explanations in science. This project will be supervised by Dr. David Brown, a professor in Science Education at the Department of Curriculum and Instruction.

I am interested in ways to help students develop the abilities and skills needed to construct and assess explanations in science. Such an instructional outcome is an indispensable goal of science. However, research in science education has shown that students face difficulties in constructing explanations in science. Therefore, this study is aimed at helping researchers find better ways to improve students' construction of scientific explanations.

Taking part in this project entails participating in two separate individual interviews about everyday science phenomena. As experts in science, your answers to the questions will form a reference benchmark to what an expert scientific explanation looks like. This will help us in further examining students' scientific explanations to the same phenomena.

Each interview is expected to last between 40 to 50 minutes. I will conduct the interviews, which will be videotaped with your permission. I will work with you to schedule the interviews at times that are most convenient to you.

If you participate in this study you will receive $20 cash value after completing the second interview.

Your participation in this project in completely voluntary, and you are free to withdraw at any time and for any reason without penalty. You are also free to refuse to answer any questions you do not wish to answer. Your decision to participate or not participate in this project will have no consequences and will not affect your relationship with the University of Illinois in any shape or form. Any information collected during the study will be kept in strict confidence. The results of the research will be presented in professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excerpts to us.

You will get to sign an informed consent form that details and ensures all your rights as a participant. If you wish, you can receive a copy or an executive summary of the research results after the project is completed.

If you have any questions about this research project, please contact Mrs. Sahar Alameh by telephone at 217-979-5471 (call or text) OR by email at alameh2@illinois.edu OR Dr. David Brown at debrown@illinois.edu. Their address location is the following:

Department of Curriculum and Instruction
University of Illinois at Urbana-Champaign
Education Building, 1310 S Sixth Street
Champaign, Illinois, 61820

I am hopeful you will agree to help with this research project.

Best Regard,

Sahar Alameh

College of Education
Curriculum & Instruction

IRB Number: 19606

Dr. David Brown
Mrs. Sahar Alameh

Office 384 Education Building
1310 S. Sixth St.
Champaign, IL 61820

INVITATION TO PARTICIPATE IN A RESEARCH STUDY:

The Nature of Scientific Explanation (NOSE): Using a Philosophically Guided Framework to Examine the Nature and Quality of Scientific Explanations Constructed by Freshmen College Students, Science Teachers, and Practicing Scientists

WHAT IS THE STUDY ABOUT?

In this study we are interested in studying how participants construct their explanations of scientific phenomena.

WHO CAN PARTICIPATE?

- College Freshmen Students (must have completed at least two years of high school science)
- High School Science Teachers
- Scientists & Science Professors
- Science PhD Candidates (Advanced stage – ABD/dissertation stage)
- Science Postdoc Fellows

EACH PARTICIPANT WILL GET $20 CASH FOR PARTICIPATING IN THE STUDY!

WHAT IS INVOLVED?

Two interviews that will take 40-50 minutes at a time convenient to you. Interviews will take place on University of Illinois at Urbana Champaign Campus.

WHAT ARE THE BENEFITS?

You will help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

PARTICIPATION IS VOLUNTARY AND PERSONAL INFORMATION WILL BE KEPT CONFIDENTIAL!

If you have any questions or are interested in participating in the study: Text or Call Sahar Alameh on 217-979-5471 or email at alameh2@illinois.edu

Are you a high school science teacher?

Do you want to participate in a research study?

- **Two fun science-based interviews**
- **Interviews about 40-50 minutes each**
- **You receive $20 for participating**

- **If interested, or if you have any questions please email Sahar Alameh at alameh2@illinois.edu**

IRB Number: 19606

I ILLINOIS | College of Education

Are you a freshman student with high-school science background?

Do you want to participate in a research study?

- **Two fun science-based interviews**
- **Interviews about 40-50 minutes each**
- **You receive $20 for participating**

- **If interested, or if you have any questions please email Sahar Alameh at alameh2@illinois.edu**

IRB Number: 19606

I ILLINOIS | College of Education

Are you a science doctoral candidate (ABD/advanced PhD stage)?

Do you want to participate in a research study?

- **Two fun science-based interviews**
- **Interviews about 40-50 minutes each**
- **You receive $20 for participating**

- **If interested, or if you have any questions please email Sahar Alameh at alameh2@illinois.edu**

IRB Number: 19606

I ILLINOIS | College of Education

APPENDIX B

CONSENT LETTERS

University of Illinois at Urbana-Champaign
Institutional Review Board # 19606

Consent Form for the Participant Freshmen College Student

You are being asked to participate in a voluntary research study. The purpose of this study is to explore students' explanations and their views of the *goodness* of explanations within the context of science. We are interested in ways to help students develop the abilities and skills needed to construct and assess explanations in science. I hope that the participation in this research may benefits you personally. But even if it does not, study findings are anticipated to help students develop the abilities and skills needed to construct and assess explanations in science.

If you agree to take part in this study, you will be asked to participate in two individual interviews. Each interview will be videotaped and is expected to last between 40 and 50 minutes. In the first interview, you will be asked to explain various phenomena in science. In the second interview, you will be asked to examine and comment on a number of explanations related to the science phenomena. In particular, you will listen to recordings of other individuals' explanations about the same phenomena. Recordings of your own explanations might be used in these interviews. However, any audio segments that are related to your identity, profession, or any personal information will be deleted from the excerpt. Risks and discomfort related to this research are not different from those associated with everyday life. Potential risks or discomforts include mild anxiety felt while attempting to answer interview questions. The researcher will attempt to make the interviews as comfortable as possible for you. Your participation in this research may help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

Principal Investigator: David Brown, PhD
Department and Institution: Curriculum and Instruction, University of Illinois and Urbana Champaign
Contact Information: debrown@illinois.edu

Researcher: Sahar Alameh, Doctoral Candidate
Department and Institution: Curriculum and Instruction, University of Illinois and Urbana Champaign
Contact Information: alameh2@illinois.edu

Why am I being asked?
You have been asked to participate in this research because **you are a Freshman college student with a background in science.** Approximately 30 participants will be involved in this research – approximately 10 of which are Freshmen college students at the University of Illinois at Urbana-Champaign.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois at Urbana-Champaign. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

What procedures are involved?
If you agree to take part in this study, you will be asked to participate in two individual interviews. Each interview will be videotaped and is expected to last between 40 and 50 minutes. In the first interview, you will be asked to explain various phenomena in science. In the second interview, you will be asked to examine and comment on a number of explanations related to the science phenomena. In particular, you

will listen to recordings of other individuals' explanations about the same phenomena. Recordings of your own explanations might be used in these interviews. However, any audio segments that are related to your identity, profession, or any personal information will be deleted from the excerpt. I will work with you to schedule the interviews at times that are most convenient to you and I am ready to conduct the interview in a place of your choosing.

What are the potential risks and discomforts?
Risks and discomfort related to this research are not different from those associated with everyday life. Potential risks or discomforts include mild anxiety felt while attempting to answer interview questions. The researcher will attempt to make the interviews as comfortable as possible for you.

Are there benefits to participating in the research?
Your participation may help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

What other options are there?
You have the option to not participate in this study.

Will my study-related information be kept confidential?
The information gathered during this study will remain confidential during this project. Faculty, students, and staff who may see your information will maintain confidentiality to the extent of laws and university policies. During the second interview, participants will listen to audio excerpts of one another in order to examine participants' perceptions about scientific explanations. Parts of your interview might be chosen for that purpose. However, we will remove all identifying information including but not limited to your name, age, job, etc. Our focus strictly pertains to the content and ideas you provide us. There will not be any identifying names on the tapes, and your name will not be available to anyone. The results of the research will be presented professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excerpts to us.

Will I be reimbursed for any expenses or paid for my participation in this research?
At the completion of the second interview, each participant will be offered a cash payment of $20 for his/her participation in this research.

Can I withdraw or be removed from the study?
If you decide to participate, you are free to withdraw your consent and discontinue participation at any time. The researchers also have the right to stop your participation in this study without your consent if they believe it is in your best interests, you were to object to any future changes that may be made in the study plan.

Will data collected from me be used for any other research?
Your de-identified name could be used for future research without additional informed consent. In addition, de-identified information will not include recordings even after names are removed.
The results of the research will be presented professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excepts to us.

Do you allow for the research team to use your audio excerpts when disseminating results? __Yes __No. Please note that we will obtain permission for the specific excerpt before sharing it

Who should I contact if I have questions?
Contact the researchers Dr. David Brown at debrown@illinois.edu OR Sahar Alameh at alameh2@illinois.edu if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

What are my rights as a research subject?
If you have any questions about your rights as a participant in this study, please contact the University of Illinois at Urbana-Champaign Office for the Protection of Research Subjects at 217-333-2670 or irb@illinois.edu.

I have read the above information. I have been given an opportunity to ask questions and my questions have been answered to my satisfaction. I agree to participate in this research. I will be given a copy of this signed and dated form.

________________________________	_____________
Signature	Date
________________________________	
Printed Name	
________________________________	_____________
Signature of Person Obtaining Consent	Date (must be same as subject's)
________________________________	
Printed Name of Person Obtaining Consent	

University of Illinois at Urbana-Champaign
Institutional Review Board # 19606

Consent Form for the Participant Secondary Science Teacher

You are being asked to participate in a voluntary research study. The purpose of this study is to explore students' explanations and their views of the *goodness* of explanations within the context of science. We are interested in ways to help students develop the abilities and skills needed to construct and assess explanations in science. The interviews are not meant to evaluate your answers, but rather learn about the ways by which science teachers explain phenomena and what they think about explanations. This will greatly help us gain a better understanding for examining students' scientific explanations.
If you agree to take part in this study, you will be asked to participate in two individual interviews. Each interview will be videotaped and is expected to last between 40 and 50 minutes. In the first interview, you will be asked to explain various phenomena in science. In the second interview, you will be asked to examine and comment on a number of explanations related to the science phenomena. In particular, you will listen to recordings of other individuals' explanations about the same phenomena. Recordings of your own explanations might be used in these interviews. However, any audio segments that are related to your identity, profession, or any personal information will be deleted from the excerpt. Risks and discomfort related to this research are not different from those associated with everyday life. Potential risks or discomforts include mild anxiety felt while attempting to answer interview questions. The researcher will attempt to make the interviews as comfortable as possible for you. Your participation may help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

Principal Investigator Name and Title: David Brown, PhD
Department and Institution: Curriculum and Instruction, University of Illinois and Urbana Champaign
Contact Information: debrown@illinois.edu

Researcher: Sahar Alameh, Doctoral Candidate
Department and Institution: Curriculum and Instruction, University of Illinois and Urbana Champaign
Contact Information: alameh2@illinois.edu

Why am I being asked?
You have been asked to participate in this research because **you are a secondary science teacher in Illinois.** Approximately 30 participants will be involved in this research – 10 of which are secondary science teachers in Illinois.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois at Urbana-Champaign. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

What procedures are involved?
If you agree to take part in this study, you will be asked to participate in two individual interviews. Each interview will be videotaped and is expected to last between 40 and 50 minutes. In the first interview, you will be asked to explain various phenomena in science. In the second interview, you will be asked to

examine and comment on a number of explanations related to the science phenomena. In particular, you will listen to recordings of other individuals' explanations about the same phenomena. Recordings of your own explanations might be used in these interviews. However, any audio segments that are related to your identity, profession, or any personal information will be deleted from the excerpt. I will work with you to schedule the interviews at times that are most convenient to you and I am ready to conduct the interview in a place of your choosing.

What are the potential risks and discomforts?
Risks and discomfort related to this research are not different from those associated with everyday life. Potential risks or discomforts include mild anxiety felt while attempting to answer interview questions. The researcher will attempt to make the interviews as comfortable as possible for you.

Are there benefits to participating in the research?
Your participation may help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

What other options are there?
You have the option to not participate in this study.

Will my study-related information be kept confidential?
The information gathered during this study will remain confidential during this project. Faculty, students, and staff who may see your information will maintain confidentiality to the extent of laws and university policies. During the second interview, participants will listen to audio excerpts of one another in order to examine participants' perceptions about scientific explanations. Parts of your interview might be chosen for that purpose. However, we will remove all identifying information including but not limited to your name, age, job, etc. Our focus strictly pertains to the content and ideas you provide us. There will not be any identifying names on the tapes, and your name will not be available to anyone. The results of the research will be presented professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excerpts to us.

Will I be reimbursed for any expenses or paid for my participation in this research?
At the completion of the second interview, each participant will be offered a payment of $20 cash value for his/her participation in this research.

Can I withdraw or be removed from the study?
If you decide to participate, you are free to withdraw your consent and discontinue participation at any time. The researchers also have the right to stop your participation in this study without your consent if they believe it is in your best interests, you were to object to any future changes that may be made in the study plan.

Will data collected from me be used for any other research?
Your de-identified name could be used for future research without additional informed consent. In addition, de-identified information will not include recordings even after names are removed. The results of the research will be presented professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excepts to us.

Do you allow for the research team to use your audio excerpts when disseminating results? __Yes __No. Please note that we will obtain permission for the specific excerpt before sharing it

Who should I contact if I have questions?
Contact the researchers Dr. David Brown at debrown@illinois.edu OR Sahar Alameh at alameh2@illinois.edu if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

What are my rights as a research subject?
If you have any questions about your rights as a participant in this study, please contact the University of Illinois at Urbana-Champaign Office for the Protection of Research Subjects at 217-333-2670 or irb@illinois.edu.

I have read the above information. I have been given an opportunity to ask questions and my questions have been answered to my satisfaction. I agree to participate in this research. I will be given a copy of this signed and dated form.

__________________________________ ______________
Signature Date

____
Printed Name

__________________________________ ______________
Signature of Person Obtaining Consent Date (must be same as subject's)

____
Printed Name of Person Obtaining Consent

University of Illinois at Urbana-Champaign
Institutional Review Board # 19606

Consent Form for the Participant Practicing Scientist

You are being asked to participate in a voluntary research study. The purpose of this study is to explore students' explanations and their views of the *goodness* of explanations within the context of science. We are interested in ways to help students develop the abilities and skills needed to construct and assess explanations in science. As an expert in science, your answers to the questions will form a reference benchmark to what an expert scientific explanation looks like. This will help us in further examining students' scientific explanations to the same phenomena.

If you agree to take part in this study, you will be asked to participate in two individual interviews. Each interview will be videotaped and is expected to last between 40 and 50 minutes. In the first interview, you will be asked to explain various phenomena in science. In the second interview, you will be asked to examine and comment on a number of explanations related to the science phenomena. In particular, you will listen to recordings of other individuals' explanations about the same phenomena. Recordings of your own explanations might be used in these interviews. However, any audio segments that are related to your identity, profession, or any personal information will be deleted from the excerpt. Risks and discomfort related to this research are not different from those associated with everyday life. Potential risks or discomforts include mild anxiety felt while attempting to answer interview questions. The researcher will attempt to make the interviews as comfortable as possible for you. Your participation may help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

Principal Investigator Name and Title: David Brown, PhD
Department and Institution: Curriculum and Instruction, University of Illinois and Urbana Champaign
Contact Information: debrown@illinois.edu

Researcher: Sahar Alameh, Doctoral Candidate
Department and Institution: Curriculum and Instruction, University of Illinois and Urbana Champaign
Contact Information: alameh2@illinois.edu

Why am I being asked?
You have been asked to participate in this research because **you are a practicing scientist (including graduate students in the final stage of doctoral program, postdoctoral fellows, or professional scientists).** Approximately 30 participants will be involved in this research – approximately 10 of which are practicing scientist.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois at Urbana-Champaign. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

What procedures are involved?
If you agree to take part in this study, you will be asked to participate in two individual interviews. Each interview will be videotaped and is expected to last between 40 and 50 minutes. In the first interview, you will be asked to explain various phenomena in science. In the second interview, you will be asked to examine and comment on a number of explanations related to the science phenomena. In particular, you will listen to recordings of other individuals' explanations about the same phenomena. Recordings of your own explanations might be used in these interviews. However, any audio segments that are related to your identity, profession, or any personal information will be deleted from the excerpt.

I will work with you to schedule the interviews at times that are most convenient to you.

What are the potential risks and discomforts?
Risks and discomfort related to this research are not different from those associated with everyday life. Potential risks or discomforts include mild anxiety felt while attempting to answer interview questions. The researcher will attempt to make the interviews as comfortable as possible for you.

Are there benefits to participating in the research?
Your participation may help science educators to develop instruction that helps high school and Freshman college students develop good explanations and attain conceptual understanding in science.

What other options are there?
You have the option to not participate in this study.

Will my study-related information be kept confidential?
The information gathered during this study will remain confidential during this project. Faculty, students, and staff who may see your information will maintain confidentiality to the extent of laws and university policies. During the second interview, participants will listen to audio excerpts of one another in order to examine participants' perceptions about scientific explanations. Parts of your interview might be chosen for that purpose. However, we will remove all identifying information including but not limited to your name, age, job, etc. Our focus strictly pertains to the content and ideas you provide us. There will not be any identifying names on the tapes, and your name will not be available to anyone. The results of the research will be presented professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excerpts to us.

Will I be reimbursed for any expenses or paid for my participation in this research?
At the completion of the second interview, each participant will be offered a payment of $20 cash value for his/her participation in this research.

Can I withdraw or be removed from the study?
If you decide to participate, you are free to withdraw your consent and discontinue participation at any time. The researchers also have the right to stop your participation in this study without your consent if they believe it is in your best interests, you were to object to any future changes that may be made in the study plan.

Will data collected from me be used for any other research?
Your de-identified name could be used for future research without additional informed consent. In addition, de-identified information will not include recordings even after names are removed. The results of the research will be presented professional meetings and published in scholarly journals. However, data will be reported in aggregate form and will not include any identifying information. Pseudonyms will be used, and identifying information (such as names) will be deleted, in case quotes or audio excerpts are used in reports. In case we decide to use audio excerpts from your interviews when disseminating the results of the project, we will first seek your active consent to release such excepts to us.

Do you allow for the research team to use your audio excerpts when disseminating results? __Yes __No. Please note that we will obtain permission for the specific excerpt before sharing it

Who should I contact if I have questions?

Contact the researchers Dr. David Brown at debrown@illinois.edu OR Sahar Alameh at alameh2@illinois.edu if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

What are my rights as a research subject?
If you have any questions about your rights as a participant in this study, please contact the University of Illinois at Urbana-Champaign Office for the Protection of Research Subjects at 217-333-2670 or irb@illinois.edu.

I have read the above information. I have been given an opportunity to ask questions and my questions have been answered to my satisfaction. I agree to participate in this research. I will be given a copy of this signed and dated form.

__________________________________ ______________
Signature Date

____
Printed Name

__________________________________ ______________
Signature of Person Obtaining Consent Date (must be same as subject's)

____
Printed Name of Person Obtaining Consent

APPENDIX C

PROTOCOL FOR INTERVIEW I

Introduction

Thank you for participating in this interview. I would like to get some background information about you before we start. This information will be held in utmost confidentiality and will only be accessible to me and my advisors. In this interview, I will ask you about different topics in science. I am interested in knowing more about your responses to these topics. Please feel free to express what is on your mind as there is no right or wrong answers to any question I am going to ask. My goal is not to evaluate your answers. Do you have any questions for me before we begin?

Personal Information (All participants)

Code:____________________

Sex:_____________________

Age:_____________________

Ethnicity (Optional):_____________________

Time interview began:____________________

Time interview ended:____________________

Contact Information:

(a) Email:______________________________

(b) Phone Number:______________________

Education and professional background:

Freshmen students ONLY:

What are the high school science courses you have completed?______________________________________

____

Have you taken any AP courses? If yes, what are they?____

____

What is your college major?

How would you rate your achievement in science in high school on a scale from 1 (poor) to 5 (excellent)?

How would you rate your understanding of science and general science concepts on a scale from 1 (poor) to 5 (excellent)?

Are there any other outstanding experiences related to science learning, science teaching or practicing science that you would like to share with me?

Secondary science teachers ONLY:

What is your highest degree?

What is your undergraduate college major?

What is your undergraduate college minor?

How many years of teaching experience do you have?

What are the level(s) you have taught?

What content area(s) have you taught?

Are there any other outstanding experiences related to science learning, science teaching or practicing science that you would like to share with me?____

____

Scientists ONLY:

What is your highest degree?____

____

What year was it granted (or expected to be granted)?___________________________________

____

What is/are your field(s) of expertise?____

____

For doctoral students: At what stage of your doctoral program are you currently in?________________

____

How do you describe your major research interests in lay terms?____________________________

____

Are there any other outstanding experiences related to science learning, science teaching or practicing science that you would like to share with me?____

____

The interview will now begin. Do I have your consent to videotape this interview?

Just a heads up, throughout this interview, I will always ask you questions such as “Is there anything you would like to add to your explanation to make it complete?” These questions will be asked whether or not I think your answers or explanations are complete or good. I will keep asking it until you tell me you don’t have anything else to add. Remember, this interview does not aim to evaluate your answers. There are no correct/wrong answers.

Interview Scenario I: ***(The Dancing Raisins)***

In this activity you will observe a phenomenon using the following materials I have on this table (*Interviewer points out to the 7UP bottle, clear glass, and several fresh raisins).* You will be asked to predict what will happen and provide an explanation for your observations.

1. What do you think will happen when I place the raisins in a glass of 7-UP? Why?
    a. Will they sink or float?
    b. Why will the raisins sink/float?
    c. Is there anything you would like to, can or should add to make your explanation complete?
    d. When participant decides that it is complete, ask: do you think that your explanation is adequate in explaining why you think this will happen?

Let us know see what actually happens. You can record your observation on a piece of paper if you wish to do so. *The interviewer fills the glass with 7-UP and drops a few raisins into the glass.*

2. Describe what happened when I placed the raisins in the glass with 7-UP?
    a. Ask about the recorded observation, if applicable
3. Did your predictions align with your observations? Why? Why not?
4. Why do you think the raisins first sank to the bottom? Why did they then float up to the top then sink again? (*Reword based on the interviewee answer).*
5. *When raisins stop 'dancing':* Why do you think the raisins stopped sinking to the bottom and then floating up?
6. Why do you think pop tastes "flat" after it's be out for a while?
7. Is there anything you would like to, can or should add to make your explanation complete?
8. When participant decides that it is complete, ask: do you think that your explanation is adequate in explaining why you think this happens?

Probing questions:

1. What is carbonation? How can you tell it is in the 7-Up?

2. What is density?
3. Which is denser: raisins of soda pop? How can you tell?

Final question: Ok now let us wrap up. Can you provide a final explanation to this activity? In other words, can you describe what happened when I put raisins in a glass of 7UP and why you think this happened? Be as detailed and thorough as possible. Remember, I am not evaluating your answer, I am interested in the way you construct your scientific explanation.
Prior knowledge: Are you familiar with this phenomenon? Have you seen this or something similar to this phenomenon before? Where? Can you elaborate?

Interview Scenario II: ***(The Classic Candle Experiment)***

In this activity you will observe a phenomenon using the following materials I have on this table (*Interviewer points out to the candle, plate, food color, and a jar*). You will be asked to predict what will happen and provide an explanation for your observations.

- The interviewer *secures the candle on the plate using sticky putty.*
- *The interviewer pours water into the plate and add a few drops of food color.*
- *The interviewer lights* the candle.

1. What do you think will happen if I cover the candle with an upside down glass jar (point out to the jar)?
    a. Why do you think this will happen?
    b. Is there anything you would like to, can or should add to make your explanation complete?
    c. When participant decides that it is complete, ask: do you think that your explanation is adequate in explaining why you think this will happen?

    Now let us perform this step and observe what will happen. You can record your observation on a piece of paper if you wish to do so. *The interviewer now covers the candle with the upside down jar. The candle flame gradually diminishes before expiring. In addition, the water level rises very slowly (if at all) as the candle flame diminishes, and rises quickly after the flame has completely expired.*
2. Describe what happened when I covered the candle with the upside down glass jar.
    a. Ask about the recorded observation, if applicable
3. Did your observations align with your predictions? Why? Why not?

4. Why do you think this happened?
    a. Is there anything you would like to, can or should add to make your explanation complete?
    b. When participant decides that it is complete, ask: do you think that your explanation is adequate in explaining why you think this happens?

Probing questions:

1- Why did the flame diminish?
    a. Did it diminish quickly or slowly? Why do you think this happened?

2- Why did the water level rise?
    a. When did the water level rise?
    b. Why didn't the water start rising from the instant the candle is covered?
    c. Why didn't the water stop as soon as the flame expired?
    d. Did it rise quickly or slowly? Why do you think this happened?

3- If noticeable or recorded by the interviewee: How do you explain the air bubbles escaping from the jar?

4- What do you think might happen if I use a smaller jar? Why do you think this would happen?
    a. A bigger jar? Why do you think this would happen?
    b. A smaller candle? Why do you think this would happen?
    c. A bigger candle? Why do you think this would happen?

5- If participant offers alternative explanations for what happened, interview will ask about the way the participant would design a test to assess these alternative explanations (e.g., water rises because oxygen burns vs. because of the expansion of heated gas and then cooling after the candle dies).

Note: Extra material will be available if the interviewee wishes to use a bigger/small jar, and/or a bigger/smaller candle – and if time allows.

Final question: Ok now let us wrap up. Can you provide a final explanation to this activity? In other words, can you describe what happened when I covered the lit candle with an upside jar and

why you think this happened? Be as detailed and thorough as possible. Remember, I am not evaluating your answer, I am interested in the way you construct your scientific explanation.
Prior knowledge: Are you familiar with this phenomenon? Have you seen this or something similar to this phenomenon before? Where? Can you elaborate?

More Explain-Only Scenarios (Videos)

1. Watch Video: Why does the penny seem higher in the water? (Or why do we see two pennies)?

Final question: Ok now let us wrap up. Can you provide a final explanation to this activity? In other words, can you describe what you saw here and why it happened this way? Be as detailed and thorough as possible. Remember, I am not evaluating your answer, I am interested in the way you construct your scientific explanation.

Prior knowledge: Are you familiar with this phenomenon? Have you seen this or something similar to this phenomenon before? Where? Can you elaborate?

2. Watch video: Why does food coloring spread out faster in hot water than in cold water?

Final question: Ok now let us wrap up. Can you provide a final explanation to this activity? In other words, can you describe what you saw here and why it happened this way? Be as detailed and thorough as possible. Remember, I am not evaluating your answer, I am interested in the way you construct your scientific explanation.

Prior knowledge: Are you familiar with this phenomenon? Have you seen this or something similar to this phenomenon before? Where? Can you elaborate?

At all times ask:

a. Is there anything you would like to, can or should add to make your explanation complete?
b. When participant decides that it is complete, ask: do you think that your explanation is adequate in explaining why you think this happens?
c. Follow up questions and prompts as necessary.

APPENDIX D
PROTOCOL FOR INTERVIEW II

As you may recall, in the previous interview you provided some explanations to a few scientific phenomena. In this interview, we will be doing two things:

First, I am going to show you a diagram that is meant to represent one of the explanations that you have made in the previous interview. I will also show you the transcript of your explanation that corresponds to the diagram. I will ask you to look at this diagram and tell me whether or not you think it represents your explanation accurately. Of course, I do not expect that you will remember the specific details of all the explanations you provided. I am interested in your general ideas as you examine these diagrams and whether you think the diagrams are consistent with what you believe.

1- Before showing NOSE explanation diagram ask: In your opinion, what is a scientific explanation?

The interview examines his/her explanation diagrams based on NOSE and provides feedback to the researcher. The researcher will attempt to clarify these comments by asking, when applicable, probing questions, such as, "Is there any way I can change this diagram to better represent your explanation."

Second, I am going to present to you transcripts of explanations in which other individuals had responded to various phenomena that we discussed during first interview together. I will show you each transcript separately for each phenomena for each individual. You can take all the time you need to read each transcript. I would like for you to comment on the validity or adequacy of the explanations of these individuals.

The researcher shows the transcripts and asks the interviewee to comment on the explanations being presented. The following questions will be used to prompt the interviewee:

1- What do you think of the explanation proposed by this person?

2- What are the strong aspects of this explanation?

3- What are the weak aspects of this explanation?

4- Would you consider this to be an adequate, valid or good explanation? Why or why not?

5- Would you consider this to be a complete explanation? Is there anything that can be added to it to make it complete?

After the interviewee comments on all recordings per phenomena, the following will be asked:

1- What, in your own words, are the criteria that you used to assess the adequacy of the explanations that you listened to?

2- What, in your own words, are the criteria that you used to assess the completeness of the explanations that you listened to?

3- I would like to ask you again, in your opinion what is a scientific explanation?