Cause and Because: Using Epistemic Network Analysis to Model Causality in the Next Generation Science Standards

Amanda Siebert-Evenstone[1](✉) and David Williamson Shaffer[1,2]

[1] University of Wisconsin – Madison, Madison, WI, USA
alevenstone@wisc.edu
[2] Aalborg University, Copenhagen, Denmark

Abstract. The Next Generation Science Standards propose an integrated and holistic view of science education that teaches science through three-dimensional learning. In this vision of science, content and practices are interconnected and inseparable. While the NGSS has influenced K-12 education standards in 40 states, there has not been a systematic analysis of the standards themselves. In this study, we investigate three-dimensional learning in order to identify new insights into underlying relationships between science concepts as well as make comparisons between different science disciplines. We used Epistemic Network Analysis to measure and models the structure of connections among crosscutting concepts and practices within and across disciplines. Results show systematic differences between how Physical and Life Sciences use and describe cause and effect relationships in which Physical Sciences predominantly focuses on the generation of causal relationships while Life Sciences focuses on the explanation of causal relationships.

Keywords: Epistemic network analysis · Next generation science standards · Three-dimensional learning

1 Introduction

Following the turn in science education toward teaching science as a practice [1], the Next Generation Science Standards (NGSS) [2, 3] constructed a practice-based vision for science education in the United States. The NGSS propose an integrated and holistic view of science education that organizes science into *three-dimensional learning*: a coherent combination of disciplinary core ideas, crosscutting concepts, and science and engineering practices. In this vision of science, content and practices are interconnected and inseparable. As such, this document provides an important artifact of what scientists and science educators deem valuable and core to the pursuits of this discipline and how students could learn how to think like scientists.

While there are many articles, books, and websites that provide resources for teacher implementation, there has been less research on the implications and rhetoric of the standards themselves. In this study, we investigate overarching claims about the

B. Eagan et al. (Eds.): ICQE 2019, CCIS 1112, pp. 223–233, 2019.
https://doi.org/10.1007/978-3-030-33232-7_19

interconnected nature of the NGSS, specifically what are the relationships among the three dimensions of science learning.

2 Theory

With the goal of improving K-12 science education, the Next Generation Science Standards were developed through a collaboration between the National Research Council (NRC), the National Science Teachers Association, the American Association for Advancement of Science, and Achieve, Inc. [2, 3]. This project united science experts, researchers, and educators to create a new vision for science education and consequently a new set of education standards to be followed in K-12 classrooms. These standards proposed and organized important and overarching themes in science into what the standards call 3-dimensional learning including

Dimension 1: *Science and Engineering Practices*, which are the skills and knowledge scientists and engineers employ;
Dimension 2: *Crosscutting Concepts*, which are the common themes and unifying ideas across the disciplines; and
Dimension 3: *Disciplinary Core Ideas*, which are specific and fundamental concepts and contexts necessary for understanding the discipline.

In this vision of science, content and practices are interconnected and inseparable. Instead of learning content and then applying it, the NGSS proposed an integrated and holistic view of science education. As such, this document provides an important artifact of what scientists and science educators deem valuable and core to the pursuits of this discipline and how students could learn how to think like scientists.

While the NGSS has influenced K-12 education standards in 40 states, there has not been a systematic analysis of the standards themselves. Recent work has analyzed components of the standards, such as genetics content [4], sustainability [5], or a single crosscutting concept (i.e. scale, proportion, and quantity) [6]. One reason there may have been few systematic analyses is that the publicly available version of the standards is an unwieldy and dense set of tables within a lengthy document.

To dive deeper into this conception of science thinking I use David Shaffer's [7] epistemic frame theory to describe the pattern of associations among skills, knowledge, and other cognitive elements that characterize groups of people who share similar ways of framing, investigating, and solving complex problems. More specifically, epistemic frame theory considers the ways in which certain groups of people think and suggests that in specific communities there is a set of systematic patterns of relationships among skills, knowledge, identity, values, and epistemology that form the epistemic frame for that community.

Importantly, epistemic frame theory shifts the focus of learning from accumulating isolated pieces of knowledge to focusing on the structure of connections among them. Similarly, diSessa [8] argued that deep understanding results from linking basic disciplinary concepts within a theoretical framework. For example, diSessa describes how novices have "knowledge-in-pieces", whereas experts have a deep and systematic understanding of how these disciplinary concepts are connected. Other learning scientists have similarly conceptualized learning as developing patterns of connections between concepts [9, 10].

Therefore, in order to model what it means to adopt the epistemic frame of a scientist, or more simply what it means to think like a scientist, we need a way to analyze the relationships among elements in that domain. One way to measure the relationships among elements in an epistemic frame is by using epistemic network analysis (ENA), a tool designed to analyze the structure of connections by identifying the co-occurrence of domain elements in a particular community of practice [11]. The resulting models can be visualized as networks in which the nodes in the model are the codes and the lines connecting the nodes represent the co-occurrence of two codes. Thus, I can quantify and visualize the structure of connections between science practices and crosscutting concepts making it possible to characterize important connections for each science discipline.

In this study, I investigate how modeling and measuring the connections between practices and concepts can identify new insights into underlying relationships between science concepts as well as make comparisons between different science disciplines.

3 Methods

3.1 Data Source

The NGSS provide a set of 208 K-12 science standards organized across three science disciplines (Earth and Space Sciences, Life Sciences, and Physical Sciences) as well as sections addressing Engineering and Technical Sciences. Each discipline has standards that are arranged by Performance Expectations (PEs) that constitute what should be learned by students by the end of that grade level.

3.2 Segmentation and Coding

In this analysis, each line of data represents a single chunk of written content from the NGSS. For example, in the performance expectation for MS-PS4-1, the standards outline 6 pieces of information and each unique piece of text was segmented into a different row. Based on the structure of the NGSS layout, each performance expectation was further segmented by the specific science and engineering practice (SEP, Table 1) and crosscutting concept (CCC, Table 2) that was identified.

For example, MS-PS4-1 asks students to describe the amplitude of waves using mathematics and computational thinking (SEP) and identify patterns (CCC).

3.3 Epistemic Network Analysis

To analyze the connections with the NGSS, I used Epistemic Network Analysis (ENA) [11, 14], which models the structure of connections among NGSS code elements. ENA measures connections by quantifying the co-occurrence of practice and concept within a defined conversation. In this case, a conversation is a collection of lines of data such that lines within a conversation are assumed to be closely related.

Table 1. List of science and engineering practice codes adapted from National Science Teaching Association (NSTA) [12].

Science and Engineering Practice (SEP)	Definition
Asking questions and defining problems	A practice of science is to ask and refine questions that lead to descriptions and explanations of how the natural and designed world works and which can be empirically tested.
Developing and using models	A practice of both science and engineering is to use and construct models as helpful tools for representing ideas and explanations.
Planning and carrying out investigations	Scientists and engineers plan and carry out investigations in the field or laboratory, working collaboratively as well as individually. Their investigations are systematic and require clarifying what counts as data and identifying variables or parameters.
Analyzing and interpreting data	Scientific investigations produce data that must be analyzed in order to derive meaning.
Using mathematics and computational thinking	In both science and engineering, mathematics and computation are fundamental tools for representing physical variables and their relationships.
Constructing explanations and designing solutions	The products of science are explanations and the products of engineering are solutions.
Engaging in argument from evidence	Argumentation is the process by which explanations and solutions are reached.
Obtaining, evaluating, and communicating information	Scientists and engineers must be able to communicate clearly and persuasively the ideas and methods they generate. Critiquing and communicating ideas individually and in groups is a critical professional activity.

For the NGSS, I defined the conversation as a single PE. For example, across MS-PS4 there are 3 total PEs. Each separate expectation has an associated SEP and CCC which would be considered in the same conversation because these two codes specifically relate to one another based on their PE. MS-PS4-1, above, would have a connection between mathematic and computational thinking to the crosscutting concept of patterns However, MS-PS4-1 would not be considered related to MS-PS4-2 because the SEP and CCC relate to a different topic.

ENA constructs a network model for each unit of analysis, showing how the codes within a conversation are connected to one another. The resulting models can be visualized as network graphs where the nodes correspond to the codes and edges reflect the relative frequency of the connection between two codes. Thus, we can quantify and visualize the structure of connections among SEP and CCC, making it possible to characterize three-dimensional learning ideas within each discipline.

Table 2. List of Science and Engineering Practice codes adapted from the NSTA [13].

Crosscutting Concepts (CCC)	Definition
Patterns	Observed patterns of forms and events guide organization and classification, and they prompt questions about relationships and the factors that influence them.
Cause and effect	Events have causes, sometimes simple, sometimes multi-faceted. A major activity of science is investigating and explaining causal relationships and the mechanisms by which they are mediated. Such mechanisms can then be tested across given contexts and used to predict and explain events in new contexts.
Scale, proportion, quantity	In considering phenomena, it is critical to recognize what is relevant at different measures of size, time, and energy and to recognize how changes in scale, proportion, or quantity affect a system's structure or performance.
Systems and system models	Defining the system under study-specifying its boundaries and making explicit a model of that system-provides tools for understanding and testing ideas that are applicable throughout science and engineering.
Energy and matter	Flows, cycles, and conservation. Tracking fluxes of energy and matter into, out of, and within systems helps one understand the systems' possibilities and limitations.
Structure and function	The way in which an object or living thing is shaped and its substructure determine many of its properties and functions.
Stability and change	For natural and built systems alike, conditions of stability and determinants of rates of change or evolution of a system are critical elements of study.

4 Results

I used ENA to measure and model connections between practices and concepts for each discipline. For this ENA model, connections were counted for each performance expectation and accumulated across disciplinary core ideas and grades to model each of the four disciplines. In this paper, I specifically compare Physical Sciences and Life Sciences.

In Fig. 1, the network graph for Physical Sciences identifies many connections across the PEs and shows a few main connections, including Cause and Effect to Explanations, Cause and Effect to Argument, and Analyze to Patterns. On the other hand, Physical sciences (Fig. 2) shows the most connections between Cause and Effect to Investigations, Explanations to Energy, and Energy to Models.

One way to consider the differences in connections is to choose a common node and then analyze the similarities and differences in how each discipline connects to this idea. In the next section, we focus on a single concept and analyze the difference in connections to this idea.

Fig. 1. Network for physical sciences (purple). Thicker lines represent more frequent connections, thinner lines represent less frequent connections (Color figure online).

4.1 Cause and Effect

Within the crosscutting concepts, Cause and Effect is the most prominent concept within all three disciplines occurring a total of 56 times. In Life Sciences, Cause and Effect is included in 31% of PEs while in Physical Sciences this concept is included in 33% of all PEs. In their networks, both Life and Physical sciences make many connections between concepts and practices and in both sets of standards, there are many connections to Cause and Effect (seen by a larger diameter node and thick lines connecting to that node).

Another way to consider the differences between disciplines is to construct a difference graph (Fig. 2). The difference graph subtracts the edge weights of the mean networks of each unit visualizing the differences in weights. Connections represented

Fig. 2. Network for life sciences (blue). Thicker lines represent more frequent connections, thinner lines represent less frequent connections (Color figure online).

by purple lines were stronger among PS standards, while connections in blue occurred proportionally more often among LS standards (Fig. 3).

As the difference network shows, Life Sciences made proportionally more connections between Cause and Effect to Constructing Explanations and Designing Solutions as well as to Engaging in Argument from Evidence. On the other hand, Physical sciences were more likely to connect this idea of causality with Analyzing and Interpreting Data as well as Planning and Carrying out Investigations. These differences indicate different treatments of the ways cause and effect are used and treated in this representation of the disciplines. Life Sciences were more likely to link causality with ways to explain causal relationships while Physical Science expectations were more likely to propose the investigation and analysis of causal relationships.

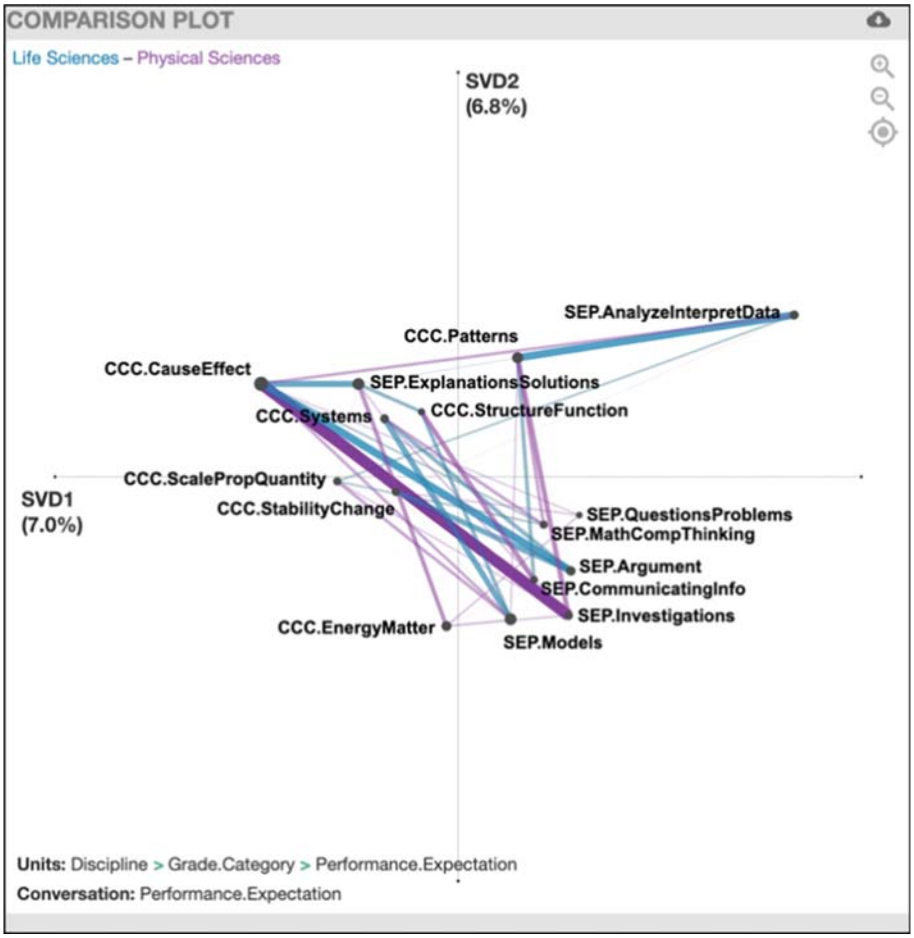

Fig. 3. Difference network between physical and life sciences, in which purple connections occur more frequently in the Physical Sciences and blue connections occur more frequently in the Life sciences (Color figure online).

This difference is also apparent in how Cause and Effect is described in the crosscutting concepts section of the performance expectations:

> "Phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability." (MS-LS1-4)

In this Life Sciences text, causality has an emphasis on the ways and potentially the only ways a certain relationship can be described which is related to both explanation and argumentation. On the other hand, Physical Sciences expectations about causality are described in a different way:

> "Simple tests can be designed to gather evidence to support or refute student ideas about causes." (1-PS4-3)

In this Physical Sciences text, there is an emphasis on using "simple tests" which is related to planning and carrying out investigations.

These are two descriptions for Cause and Effect, but there were also many cause and effect descriptions that were listed across the disciplines, including: "Events have causes that generate observable patterns" which is listed in Earth and Space Sciences, Life Sciences, and Physical Sciences. Another example is that "Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects" which can be found in both Life Sciences and Earth and Space sciences. There are even three versions of one Cause and Effect text that vary when the idea ends, "Cause and effect relationships are routinely identified [.; and used to explain change; tested, and used to explain change]". This idea and the different versions occur across the three disciplines.

However, while each of these disciplines share common language about causality, the two incline texts above were unique to their field. No other discipline listed that "phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability." Likewise, no other discipline listed that "simple tests can be designed to gather evidence to support or refute student ideas about causes." This dichotomy highlights the systematic differences with which common ideas like causality are described differently and may serve different roles across the disciplines.

5 Discussion

Preliminary evidence suggests that there are interesting and systematic differences between science disciplines. Both Life and Physical Sciences make many connections between concepts and practices and in both sets of standards, there are many connections to Cause and Effect. In comparison, Life Sciences makes more connections between causality and explanations as well as to argumentation. On the other hand, Physical Sciences more often connects this idea of causality with investigations and analysis.

Science has been defined in terms of empirically deriving causal explanations [15], and within the NGSS, they state, "A major activity of science is investigating and explaining causal relationships and the mechanisms by which they are mediated" (Appendix G, p. 1) [2]. In each of these two definitions, there are two main components, (1) generation and (2) explanation of causal relationships. Both disciplines in this analysis address these ideas, however, Physical Sciences predominantly focuses on generation while Life Sciences focuses on explanation of causal relationships. This is not to say that the physical science standards do not address explanation and argumentation or that life science standards omit investigations, rather, that in terms of causality they make more connections to the practices of investigating and analyzing data.

One goal of crosscutting concepts was to help students use ideas like cause and effect to relate and understand core ideas across disciplines. But the separation of ideas identified in this analysis may foster misconceptions for how students learn the goals of science. If the purpose of science is both investigating *and* explaining causality and students are taught *either* investigating *or* explaining these relationships students may incorrectly associate a discipline with only one component of causality. Moreover, there are important differences that should be explored in how teachers, students, and curricula unite concepts to create coherent learning experiences for students [16].

These systematic differences also have implications for how teachers and curriculum designers may create lessons and assess learning. While the NGSS and NRC *Framework* were created as guidelines for science instruction, they also serve as starting points and references for how to teach science. As such it is important to know which connections were important within and across disciplines, which can be used to compare with curricula and discussions to identify which connections occur in either, neither, or both the standards and real-world examples. Further, and more importantly, this epistemic network analysis created a metric space that can be used to measure and differentiate science thinking in other datasets. Therefore, I can create a metric space based on what the NGSS proposed and then project coded data about three-dimensional learning from real-world implementations into that space.

Of course, this analysis was limited to a comparison of two disciplines. Future analyses will further explore three-dimensional learning for earth and space sciences as well as engineering and technical sciences. Another limitation is that this analysis focuses solely on the standards themselves and does not look at science education in real classrooms. Future work will investigate how the practices and concepts are connected during actual student discussions about science and compare connection-making in the real-world with connections represented in the NGSS.

This work empirically identifies the underlying structure of the NGSS and provides both a way to compare science ideas within the standards as well as compare real-world data using the structure of the standards.

Acknowledgements. This work was funded in part by the National Science Foundation (DRL-1661036, DRL-1713110), the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

References

1. Ford, M.J., Forman, E.A.: Redefining disciplinary learning in classroom contexts. Rev. Res. Educ. **30**, 1–32 (2006). https://doi.org/10.3102/0091732X030001001
2. NGSS Lead States: Next Generation Science Standards: For States, by States (Appendix F – Science and Engineering Practices). Achieve, Inc. behalf twenty-six states partners that Collab. NGSS, pp. 1–103 (2013). https://doi.org/10.17226/18290
3. National Research Council: A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. National Academies Press, Washington, DC (2012). https://doi.org/10.17226/13165
4. Lontok, K.S., Zhang, H., Dougherty, M.J.: Assessing the genetics content in the next generation science standards. PLoS One **10**, 1–16 (2015). https://doi.org/10.1371/journal.pone.0132742

5. Feinstein, N.W., Kirchgasler, K.L.: Sustainability in science education? how the next generation science standards approach sustainability, and why it matters. Sci. Educ. **99**, 121–144 (2015). https://doi.org/10.1002/sce.21137
6. Chesnutt, K., et al.: Next generation crosscutting themes: factors that contribute to students' understandings of size and scale. J. Res. Sci. Teach. **55**(6), 876–900 (2018). https://doi.org/10.1002/tea.21443
7. Shaffer, D.W.: How Computer Games Help Children Learn. Palgrave Macmillan, New York (2006). https://doi.org/10.1057/9780230601994
8. DiSessa, A.: Knowledge in pieces. Constr. Comput. age. 49–70 (1988). https://doi.org/10.1159/000342945
9. Bransford, J.D., Brown, A.L., Cocking, R.R.: How people learn: brain, mind, experience, and school. National Academies Press, Washington D.C. (1999). https://doi.org/10.17226/9853
10. Chi, M.T.H., Feltovich, P.J., Glaser, R.: Categorization and representation of physics problems by experts and novices. Cogn. Sci. **5**, 121–152 (1981). https://doi.org/10.1207/s15516709cog0502_2
11. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison (2017)
12. National Science Teaching Association: Science and Engineering Practices (2014)
13. National Science Teaching Association: Crosscutting concepts (2014)
14. Shaffer, D.W., Ruis, A.R.: Epistemic network analysis: a worked example of theory-based learning analytics. In: Handbook of Learning Analytics Education data Mining (2017). in press. https://doi.org/10.18608/hla17.015
15. Berland, L.K., Schwarz, C.V., Krist, C., Kenyon, L., Lo, A.S., Reiser, B.J.: Epistemologies in practice: making scientific practices meaningful for students. J. Res. Sci. Teach. **53**, 1082–1112 (2016). https://doi.org/10.1002/tea.21257
16. Reiser, B.J., Mcgill, T.A.W.: Coherence from the students' perspective: why the vision of the framework for K-12 Science Requires More than Simply " Combining " Three Dimensions of Science Learning, 1 (2017)