EXPLORING AN EXPANDED CONCEPTION OF EPISTEMIC COGNITION

By

LUKE ANDREW BUCKLAND

A dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Education

Written under the direction of

Clark A. Chinn

And approved by

____

____

____

____

New Brunswick, New Jersey

October, 2015

ABSTRACT OF THE DISSERTATION

Exploring an Expanded Conception of Epistemic Cognition

by LUKE ANDREW BUCKLAND

Dissertation Director:

Clark A. Chinn

A series of three independent articles explore how the expanded framework for modeling epistemic cognition developed in Chinn, Buckland and Samarapungavan (2011) can advance psychological research. The first study: “Epistemic cognition and understanding the nature of science” investigates relations between the expanded epistemic framework and research into conceptions of the Nature of Science (NOS). This study surveys the conceptions addressed over the history of NOS research, developing a comprehensive analytic framework to trace the scope of the topics that feature in a diverse range of 81 NOS instruments used in six decades of research. By tracing historical change in the kinds of conceptions targeted for investigation, the study reveals how debate about measures and norms has led to new and modified instrumentation. The second study: “Epistemic cognition and reliable processes of knowledge production” investigates an under-researched component of the Chinn et al. (2011) framework - beliefs about the reliable processes by which knowledge is achieved. Interview and written data served to trace the epistemic beliefs about reliable processes of 19 participating undergraduates, as they reasoned about a diverse array of knowledge-generating processes. The data reveals considerable variation amongst participants in the range and kinds of processes and conditions that they considered relevant to the production of knowledge. The third study: “Epistemic growth in model-based reasoning” explores the epistemic criteria implicit in the model-based reasoning of 24 seventh-grade students engaged in inquiry learning. It conducts a fine-grained investigation of the justificatory practices of learners drawn from the classes of four teachers in two dissimilar schools, over a full school year. Data sources comprise participants' written justifications of constructed, peer, group and given models, and their comparative evaluations during model choice. Participating students adopted a complex array of higher-level criteria in their judgments about model quality, and the study provides pedagogically valuable insight into the characteristic strengths and weaknesses of their justificatory practices. Through these three inter-related studies, we argue that the expanded framework represents a viable and productive tool for advancing the field of research into epistemic cognition.

Keywords: epistemic cognition, personal epistemology, nature of science, reliable processes of knowledge production, models, modeling, model-based reasoning, science inquiry learning.

ACKNOWLEDGEMENTS

NSF grant numbers 0529582 and 1008634, on which Dr Clark. C. Chinn and Dr Ravit Golan Duncan were principal investigators, funded some of the work in this dissertation, particularly Study 3. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Thanks are due to my advisor, mentor and collaborator Clark Chinn. Thanks are also due my committee members Ravit Golan Duncan, Ala Samarapungavan, and Cindy Hmelo-Silver.

Also thanks to the Rutgers Graduate School of Education and a Rutgers Alumni scholarship in 2011 for supporting this research.

Thanks to the PRACCIS research team, including Paige Frenkel, Amanda Corigliano and Angela Goff for persevering with the search, organization, and analysis of data.

Special thanks to my family and friends for their love and support.

Please address correspondence concerning this dissertation to lukebuckland@gmail.com.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

In the article: ‘Expanding the Dimensions of Epistemic Cognition: Arguments from Philosophy and Psychology’, Chinn, Buckland and Samarapungavan (2011) proposed and defended an expanded framework for conceptualizing epistemic cognition for purposes of psychological research. They argued that philosophical research in contemporary epistemology provides a rich resource for the construction of more sophisticated, fine-grained and predictive psychological models of epistemic cognition. This dissertation further explores three of the ways in which the framework can advance educational and psychological research.

The three inter-linked studies that comprise this dissertation are “Epistemic cognition and understanding the nature of science,” “Epistemic cognition and reliable processes of knowledge production,” and “Epistemic growth in model-based reasoning.” Each of these three studies presents an independent application of some aspect of the expanded epistemic framework developed in Chinn et al. (2011), and together they demonstrate the fruitfulness and utility of the framework for purposes of psychological research.

The Introduction describes the broader field of research into epistemic cognition by outlining some key features as well as important limitations of existing research. It then describes and defends the Chinn et al. (2011) framework, articulating the ways in which it contributes to and extends current conceptions in the field. Third, each of the three studies is briefly outlined, showing how each links to the expanded framework and how each aims to advance the study of epistemic cognition. Chapters I, II and III present the three studies themselves, including separate sections reviewing relevant literatures, detailing study methods, and discussing results and conclusions. Finally, the Conclusion argues that the three studies together demonstrate the viability of the Chinn et al. framework for investigating the character and variation of epistemic features of human cognition.

Epistemic Cognition

The expanded framework for epistemic cognition that is the subject of the following three studies represents an effort to extend an active body of existing research into epistemic features of human cognition. This research focuses on various research targets including *epistemological* *beliefs*, *epistemic beliefs*, *personal epistemology*, *epistemic positions*, *epistemological reflection*, *reflective judgment*, and *epistemic cognition* (EC). The body of research shares a focus on human cognition related to epistemic concepts like knowledge, justification, truth, certainty and understanding. Early work in the area includes Perry's (1968) investigation of undergraduate reasoning about knowledge, with further groundbreaking epistemological frameworks and associated studies being subsequently developed by King and Kitchener (1994, 2004), Schommer (1990), and Hofer and Pintrich (1997, 2002). These and allied researchers have investigated a range of beliefs, attitudes and practices related to conceptions of the nature and sources of knowledge. They have also investigated beliefs about the structure of knowledge, the role of justification, and people's commitments regarding the limits, certainty and value of different kinds of epistemic achievements like knowledge, truth and understanding.

Research into epistemic features of cognition has pursued at least two distinct strands of inquiry, one treating EC as a multidimensional set of beliefs and one pursuing a developmental program that traces change in EC over time. The latter program investigates developmental change in epistemic conceptions and practices during cognitive maturation. For example, one tradition (Kuhn & Weinstock, 2002) proposed that achieving epistemic sophistication involves a stage-like progression, from an early naive realism, through intermediate absolutist and relativist phases, to the final, epistemically mature "evaluativist" stage.

In contrast to the developmental program, the multidimensional program in EC research uses questionnaires and factor analysis to seek out correlations between academic and other variables, with measures developed from multidimensional EC models. Researchers have used these correlations to explain and predict variation in learning outcomes and processes. For example, Hofer and Pintrich (1997) developed a multi-dimensional EC model with four components, involving cognitions related to the certainty of knowledge, the structure of knowledge (e.g. simplicity versus complexity), the sources of knowing (e.g. experience, reasoning, authority), and the justification of knowing (e.g. justifications from self or other). In spite of the differences between their approaches, both traditions aim to provide a theoretical account of what distinguishes epistemic sophistication from naiveté and epistemic experts from novices, as well as to develop the empirical tools to detect and measure these differences.

The field of EC research has made a variety of fascinating discoveries. For example, Qian and Alvermann (1995) found that students who regarded knowledge as complex and uncertain were more likely to restructure their beliefs during a conceptual change task. Kardash and Scholes (1996) also found that belief in the certainty of knowledge was associated with a disinclination to draw conclusions on controversial questions involving conflicting evidence. In general, it appears that those who regard knowledge as complex, uncertain and sourced in experience, rather than in authority, outperform those with contrasting beliefs on a variety of learning tasks.

In spite of the richness of the field of EC research, serious challenges to the scope and findings of the field remain unanswered. For one, though EC researchers have drawn on the conceptual resources of the philosophical field of epistemology, the profusion of work in contemporary epistemology (especially in virtue and reliabilist epistemologies) represents a significant untapped resource for EC models and measures. As argued in Chinn et al. (2011) contemporary epistemologists have explored a far wider array of epistemic issues and concepts than has been recognized by EC researchers. Another critique is that EC measures tend to explain relatively little of the variation that people demonstrate in learning outcomes and practices. Measures of EC have also often tended to rely on questionnaire assessments of explicit, reflective beliefs (e.g. “What is knowledge?”), a practice that is likely to overstate the importance of general beliefs that people are capable of accessing and expressly articulating (Sandoval, 2005). EC research is thus likely to benefit from a clearer focus on both explicit and tacit epistemic beliefs, including on the dispositions and aims relevant to epistemic sophistication.

Other challenges pertain to the measures and models of EC research. For example, some researchers regard a commitment to relativism as a necessary stage in the achievement of epistemic competence, a perspective that has been characteristic of many EC frameworks since at least Perry’s groundbreaking studies of undergraduates. However, these frameworks typically do not distinguish between significantly distinct forms of relativism. For example, ontological relativism is the view that there are no determinate, human-independent matters of fact that fix the truth-value of our claims about reality. For this kind of relativist, claims about reality can only be true relative to some person or group (e.g. relative to culture for cultural relativists, relative to individuals for subjectivists). In contrast, epistemological relativism is the view that there is no independent basis for the *justification* of our claims about reality. On this view, there are no objective criteria for what counts as a good reason, good evidence or a good argument. The failure to recognize this distinction means that extant EC frameworks typically do not properly distinguish those who deny the existence of universal truths about reality from those who instead deny that there are objective facts about what beliefs are more or less well-justified (see also Greene, Azevedo & Torney-Purta, 2008).

In sum, while EC research represents a promising line of inquiry, many of the models, measures and assumptions of the field are subject to important limitations. The next section describes and motivates the EC framework proposed in Chinn et al. (2011). It further considers some of the ways the expanded framework contribute to existing conceptions of EC by overcoming many of the limitations of the models that currently dominate the field.

The Expanded Framework

The expanded EC framework presented by Chinn et al. consists of five inter-related components: (a) epistemic aims and epistemic value; (b) the structure of knowledge and other epistemic achievements; (c) the sources and justification of knowledge and other epistemic achievements, and the related epistemic stances; (d) epistemic virtues and vices; and (e) reliable and unreliable processes for achieving epistemic aims. The central goal of the expanded framework is to identify epistemic features of cognition, whether they are explicit beliefs or implicit dispositions, which can explain and predict learning processes and outcomes. Another aim is to interpret and apply the framework in a fine-grained and flexible manner when developing EC measures and models.

The Chinn et al. framework incorporates many of the conceptual components of existing EC models, yet also includes an array of important new epistemic concepts that have not been the target of much systematic investigation. Importantly, the approach addresses some of the traditional components (e.g. certainty) in new ways. Further, each component targets both individual and social manifestations of EC. Next, each of the five components of the revised model is briefly discussed, detailing in broad outline its philosophical sources as well as reasons why it is likely to be fruitful for purposes of psychology.

Epistemic aims, the first component of the framework, refer to cognitive goals associated with finding things out and engaging in inquiry. Examples include true belief, justification, knowledge, and understanding. These epistemic aims represent the targets of intellectual activity, and are the cognitive states and outcomes towards which inquirers ideally direct their efforts. The epistemic aim of knowledge has been the predominant focus of the field of epistemology since at least the time of Plato. This aim has also long dominated EC research efforts. However, contemporary philosophers have discussed a considerably wider variety of epistemic aims, including true beliefs, understanding, explanation, and wisdom (Bishop & Trout, 2005; Goldman, 1986; Haack, 1993, 2003; Kvanvig, 2003; Moser, 2002 and Zagzebski, 2009). These additional kinds of epistemic aims represent new targets for EC research. A closely related issue, one that has had an outsized role in contemporary epistemology, involves the relative *value* of different epistemic aims. EC researchers might investigate these critical yet under-researched issues by determining the degree to which people actually value and adopt various epistemic aims. They could also examine whether people adopt these aims for their own sake, or only as a means to achieving other, more practical aims. Philosophers have also distinguished between epistemic and non-epistemic aims (e.g. desire for power, social recognition, competition, etc.) and psychologists might productively investigate how these different sources of motivation interact. Thus, while almost all EC research includes reference to some epistemic aim (usually knowledge) the expanded framework conceptualizes epistemic attainments in a more variegated and fine-grained manner. It includes a greater variety of kinds of epistemic aims, the distinctions between epistemic and non-epistemic aims, as well as issues of epistemic value.

Orthodox EC models typically conceptualize the structure of knowledge in terms of a single dimension. For example, Hofer and Pintrich (1997) instantiated the structure of knowledge as a simple continuum that ranges from simple to complex. In contrast, the expanded framework treats cognitions about the structure of knowledge in terms of a multidimensional space. Following contemporary philosophy, the framework conceptualizes attainments like knowledge in a more variegated way – as universal or contextual, discrete or interconnected, probabilistic or deterministic, locally or globally coherent, personal and subjective or impersonal and objective, etc. (Code, 1991; Longino, 2002; Salmon, 1989). This allows for a more fine-grained investigation into the various ways in which learners structure and conceptualize knowledge.

The sources and justification component of the framework, incorporates a variety of existing constructs in novel ways. Contemporary epistemologists have discussed many more kinds of sources than have traditionally been considered in EC research, including perception, memory, reasoning, and testimony, as well as introspection, intuition, revelation, scriptures, special mystical or religious experiences, fiction, art and the findings of various kinds of inquiry and pseudo-inquiry (e.g. Code, 1991; Kornblith, 1985; Williams, 2002). Each of these distinct sources (and sources of justification) could be subject to new lines of EC inquiry. In addition, while contemporary work by philosophers like Coady (1992), Lackey, and Sosa (2006) have explored the role of testimony as a crucial and ubiquitous social source of justification, EC research has tended to focus more narrowly on the undifferentiated source of “authority.” Alston (2005) has also investigated more variegated features of justificatory practices than has the EC field. While almost all knowledge is the product of multiple interacting sources, EC research has tended to focus on relatively simple and stark bivalent choices, e.g. between the source of authority versus that of experience. A focus on these more nuanced interactions is likely to be a productive target of EC research in allowing for more multidimensional and fine-grained analyses.

The “epistemic stances” that feature in the expanded framework include the stance of certainty that features in extant models, but further introduces a distinction between psychological and epistemic notions of certainty. While the first involves a person’s subjective level of confidence, the second refers to claims that achieve the highest degree of justification. In contrast for EC researchers “certain” has tended to mean simply “unchanging.” The framework also distinguishes between many more of the stances that individuals might adopt with regards the epistemic status of a claim, including withholding judgment, holding an assumption as a working hypothesis, acceptance based on the simple preponderance of evidence, etc.

The considerable philosophical literature on virtue epistemology served to inspire the fourth component of the expanded framework, epistemic virtues and vices. A number of influential epistemologists have adopted this approach (e.g. Greco, 2009; Montmarquet, 1986; Zagebski, 1996). Epistemic virtues are the learned, relatively stable dispositions, the adoption of which aid in the attainment of epistemic aims. These cognitive and behavioral dispositions include open-mindedness, intellectual carefulness, perseverance, humility, vigor, flexibility, creativity, courage, thoroughness, fair-mindedness, insightfulness, impartiality, intellectual sobriety and intellectual courage. Each of these dispositions represents a way in which epistemic sophisticates might be distinguished from novices, and serve to explain the high levels of success of particular forms of inquiry. In contrast, adopting epistemic vices typically impedes the attainment of epistemic ends. Examples of vices include close-mindedness, wishful thinking, intellectual cowardice and conformity. Manifesting these vice-like dispositions can serve to explain why an individual fails to conduct successful inquiry or is epistemically unsophisticated.

As several EC researchers have recognized, people do display many of the epistemic virtues and vices identified by epistemologists. In addition, people are likely to have specific beliefs about the value and utility of these epistemic dispositions, beliefs that might affect their inquiry behavior. In general, the virtues (or beliefs about the virtues) are likely to play a significant role in explaining and predicting learning and inquiry success (e.g. Kruglanski & Webster, 1996; Nussbaum & Sinatra, 2003; Sinatra, Southerland, McConaughy & Demastes, 2003). The expanded framework contributes to the field by advocating for investigation of a wider array of virtues and vices. It also argues or the development of a more unified conception of EC, one that integrates disparate lines of inquiry on epistemic virtues into the larger body of EC research. Chinn et al. (2011) proposed that EC researchers thus investigate both the broader variety of epistemic virtues and vices that their participants exhibit, their tacit and explicit beliefs about these dispositions, and the role they play in inquiry and learning.

The final component of the expanded framework concerns cognitions related to the reliable processes used to achieve epistemic aims. Philosophers working in the field of reliabilist epistemology have investigated the causal and natural processes by which true and justified beliefs are reliably produced (e.g. Bishop & Trout, 2005; Dretske, 2000; Goldman, 1999; Kornblith, 1985; Sosa, 2001). For epistemological reliabilists, beliefs attain the status of knowledge insofar as they are both true and generated via reliable belief-forming processes. This approach thus involves a shift away from a focus on the epistemic status of the justifications that believers are able to articulate, and instead focuses on the actual reliability of the social and cognitive processes of perception, testimony and inference in producing knowledge. Another motivation for this component of the expanded framework is the extensive contemporary philosophical work on social epistemology (e.g. Coady, 1992; Goldman, 1999; Kitcher, 1993; Longino, 2002). Social epistemology addresses the ways in which social practices, norms, and institutions support and hinder the production of knowledge.

Chinn et al. (2011) have argued that cognitions related to the reliability of the processes involved in achieving epistemic aims are likely to be a productive target of EC research. Crucially, the reliability of these processes is dependent on their conditions of implementation. In general, there are conditions under which a belief-generating process is reliably truth-conducive and conditions under which it is not. For example, experiments are a reliable process for generating justified beliefs in certain contexts, but only if they meet certain preconditions (e.g. adequate control of variables, random assignment, etc.). Little EC research has focused either on belief about the reliability of processes of achieving epistemic aims, or on beliefs about the conditions upon which that reliability depends. This component of the framework thus draws attention to what are likely to be crucial features of epistemic cognition.

In sum, Chinn et al. (2011) have made a case that the expanded framework represents a more fine-grained, comprehensive, unified and flexible account of the inter-related elements of epistemic cognition. They have further maintained that the framework promises a more fruitful approach to investigating aspects of EC responsible for generating substantive variation in learning processes and outcomes. Nonetheless, the ultimate value of the framework remains an empirical question, and depends on the degree to which it can be productive for research. This dissertation thus aims to demonstrate the value of the expanded framework by implementing some of the recommendations laid out in Chinn et al. (2011) in three separate but inter-related investigations. The next section provides a brief overview of each of the three studies.

The Three Studies

The first study, "Epistemic cognition and understanding the nature of science," explores relations between the expanded epistemic framework and the extensive and well-established body of research into conceptions and understanding of the Nature of Science (NOS). The field of NOS research aims both to understand how people think about science and to find ways of improving the accuracy and utility of those conceptions, particularly for science learners. It is characterized by a vigorous and long-standing debate regarding the constructs of NOS inquiry, the right norms of science (against which to evaluate lay beliefs), and the best methods to elicit, study and teach NOS-related beliefs. However, little research has aimed to expose the kinds of conceptions targeted by NOS inquiry itself, and very few studies have surveyed more than a handful of NOS research tools (e.g. Guerra-Ramos, 2012; Lederman, 1992; Lederman et al., 1998).

The first study thus aims to contribute to the literature by helping to fill this gap. In a broad review that synthesizes a diverse range of 81 NOS assessments, the study traces historical changes in the conceptions investigated over six decades of research. The study develops a comprehensive analytic framework, using it to trace the scope of the topics and issues that feature in NOS instrumentation. By tracing historical change in the conceptions investigated, the study reveals how debate about measures and norms has led to new and modified instrumentation. The study therefore aims to provide NOS researchers with a more clearly defined set of inquiry targets concerning the range of constructs that feature in this field. It also seeks to uncover productive areas of research into people's scientific epistemologies that have received little sustained attention from the field to date.

The expanded framework for epistemic cognition defended in Chinn et al. (2011) provided the starting point for the coding scheme described in Chapter 2. In response to the analysis of NOS instrumentation, the elaboration of the five components of the expanded framework produced a fine-grained and variegated set of constructs for the current study. The resulting analytic framework charts the conceptual structure of NOS research instruments both broadly and at a high level of detail, as recommended in Chinn et al. (2011).

The second study, "Epistemic cognition and reliable processes of knowledge production," investigates a particularly under-researched component of the Chinn et al. (2011) framework – beliefs about the reliable processes through which knowledge is generated. These include causal (e.g. the physics governing the transmission of visual information), perceptual (e.g. the capacities of observers to individuate stimuli), cognitive (e.g. the constraints and affordances of memory), social (e.g. the role of testimony in transmitting true beliefs), and institutional processes (e.g. the role of citations in determining scientific status). In spite of the profusion of work on reliabilist epistemology, psychological research has not yet investigated this aspect of epistemic cognition.

In order to trace people's epistemic beliefs about the reliable processes of knowledge production, nineteen undergraduate students provided interview and written data as they engaged in guided reasoning about vignettes involving various knowledge-generating processes. As orthodox models and measures of epistemic cognition are relatively insensitive to beliefs about processes, the study developed a new set of constructs to explore this aspect of participants' thinking about knowledge. This study therefore demonstrates the utility of the expanded framework by investigating a vital but little-considered aspect of epistemic cognition.

The third study, "Epistemic growth in model-based reasoning," explores aspects of the epistemic cognition implicit in the justificatory practices of learners. Educational psychologists increasingly regard practices of scientific justification in the science classroom as a valuable resource for facilitating successful inquiry-based science learning. In addition, many consider model-based reasoning to instantiate authentic practices of inquiry, and thus to engender more epistemologically sophisticated conceptions during learning. A third strand of research has advocated for the integration of epistemic criteria for scientific reasoning into classrooms and has encouraged their use and refinement through collaborative argumentation. In spite of the promise of these three lines of inquiry, little research has studied the impact of model-based reasoning involving epistemic criteria on the justificatory practices of science learners.

This study thus extends the prior work in Pluta, Chinn and Duncan (2011) by investigating the justificatory practices of 24 participating seventh-grade students engaged in a yearlong inquiry-learning curriculum, and aims to uncover microgenetic change in the elaborative and criteria-based character of their model justifications. The study involves a fine-grained investigation of the evaluative criteria implicit in the reasoning of learners drawn from the classes of four teachers in two schools over a full school year. During their instruction, modeling activities featured in a modular life-science curriculum spanning multiple middle-school content areas, including photosynthesis, cellular membranes and genetics. Over the course of the year, students built, evaluated, justified, and successively modified an array of models in response to collaborative discourse around rich bodies of evidence and communally negotiated sets of criteria for model quality. The data sources for this study comprise participants' written justifications of constructed, peer, group and given models, and their comparative evaluations during model choice. The study reveals participants' general justificatory strategies, the role of evidence and epistemic criteria in their justifications, and the degree to which they elaborated on the relationship between models and the reasons presented in their support. Participating students exhibited a complex array of higher-level criteria in their judgments about model quality, and this study provides pedagogically valuable insight into the characteristic strengths and weaknesses of their justificatory practices, as well as the varying effects of modeling instruction.

The three inter-related but independent studies present a range of the ways in which psychological research can productively use the Chinn et al. (2011) framework. The first study uses the five categories of the expanded framework as a starting point for an expansive yet fine-grained analysis of the conceptual structure of six decades of NOS instrumentation. The second study investigates a particularly under-researched component of the expanded framework – beliefs about reliable processes of knowledge production. The third study focuses on justification, the third component of the framework. It investigates interlocking practices and concepts of justification using the notion of epistemic criteria recommended in Chinn et al. (2011). Each of these studies thus reveals the value and fruitfulness of the expanded framework for advancing the field of inquiry into the nature and character of epistemic cognition.

References

Alston, W.P. (2005). *Beyond “justification”: Dimensions of epistemic evaluation*. Ithaca, NY: Cornell University Press.

Bishop, M. A., & Trout, J. D. (2005). *Epistemology and the psychology of human judgment*. Oxford, England: Oxford University Press.

Chinn, C.A., Buckland L.A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: arguments from philosophy and psychology, *Educational Psychologist*, 46, 141–167.

Coady, C. A. J. (1992). *Testimony: A philosophical study*. Oxford, England: Oxford University Press.

Code, L. (1991). *What can she know?* Ithaca, NY: Cornell University Press.

Dretske, F. (2000). *Perception, knowledge, and belief*. Cambridge, England: Cambridge University Press.

Goldman, A. I. (1986). *Epistemology and cognition*. Cambridge, MA: Harvard University Press.

Goldman, A. I. (1999). *Knowledge in a social world*. Oxford, England: Oxford University Press.

Greco, J. (2009). Virtue epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2008 ed.). Retrieved from http://plato.stanford.edu/entries/epistemology-virtue/

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, 43(3), 142–160.

Guerra-Ramos, M. T. (2012). Teachers’ ideas about the nature of science: A critical analysis of research approaches and their contribution to pedagogical practice, *Science & Education*. 21(5), 631–655.

Haack, S. (1993). *Evidence and inquiry*. Oxford, England: Blackwell.

Haack, S. (2003). *Defending science: Between scientism and cynicism*. Amherst, NY: Prometheus Books.

Hofer, B. K. (2002). Personal epistemology as a psychological and educational construct: An introduction. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 3–14). Mahwah, NJ: Lawrence Erlbaum.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, *Review of Educational Research*, 67, 88–140.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisiting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues, *Journal of Educational Psychology*, *88*(2), 260.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.

King, P. M., & Kitchener, K. S. (2004). Judgment Model: Twenty Years of Research on Epistemic Cognition. In Hofer, B. K., & Pintrich, P. R. (Eds.). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. 37. Psychology Press.

Kitcher, P. (1993). *The advancement of science*. New York, NY: Oxford University Press.

Kornblith, H. (Ed.). (1985). *Naturalizing epistemology*. Cambridge, MA: MIT Press.

Kruglanski, A.W., & Webster, D.M. (1996). ‘Motivated closing of the mind: “Seizing” and “freezing”.’ *Psychological Review*. 103*,* 263–283.

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In Hofer, B. K., & Pintrich, P. R. (eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Erlbaum, Mahwah, NJ. 121–144.

Kvanvig, J. L. (2003). *The value of knowledge and the pursuit of understanding.* Cambridge, England: Cambridge University Press.

Lackey, J., & Sosa, E. (Eds.). (2006). *The epistemology of testimony*. Oxford, England: Oxford University Press.

Lederman, N. (1992) Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29 (4), 331–59.
Lederman, N. G., Wade, P.D., & Bell, R.L. (1998). Assessing the nature of science: What is the nature of our assessments? *Science and Education*, 7, 595–615.
Longino, H. E. (2002). *The fate of knowledge*. Princeton, NJ: Princeton University Press.
Montmarquet, J. A. (1986). Epistemic virtue. *Mind*, 96, 482–497.
Moser. P. K. (Ed.). (2002). *The Oxford Handbook of Epistemology.* Oxford, England: Oxford University Press.
Nussbaum, E. M., & Sinatra, G. M. (2003). 'Argument and conceptual engagement'. *Contemporary Educational Psychology*, 28*,* 384–395.
Perry, W. G., Jr. (1968/1999). *Forms of intellectual and ethical development in the college years: A scheme.* San Francisco, CA: Jossey-Bass.
Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, 48(5). 486–511.
Salmon, W. C. (1989). Four decades of scientific explanation. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation* (Vol. 13, pp. 3–219). Minneapolis: University of Minnesota Press.
Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, 89, 634–656.
Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498.
Sinatra, G. M., Southerland, S. A., McConaughy, F., & Demastes, J. W. (2003). 'Intentions and beliefs in students' understanding and acceptance of biological evolution'. *Journal of Research in Science Teaching*, 40, 510–528.
Sosa, E. (2001). Reliabilism and intellectual virtue. In H. Kornblith (Ed.), *Epistemology: Internalism and externalism* (pp. 147–162). Malden, MA: Blackwell.
Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology*, 87(2), 282.
Williams, B. (2002). *Truth and truthfulness.* Princeton, NJ: Princeton University Press.
Zagzebski, L. T. (1996). *Virtues of the mind*. Cambridge, England: Cambridge University Press.
Zagzebski, L. T. (2009). *On epistemology.* Belmont, CA: Wadsworth.

CHAPTER I: EPISTEMIC COGNITION AND UNDERSTANDING THE NATURE OF SCIENCE

Abstract

Since at least the 1950s, a flourishing branch of research has investigated epistemic cognition involving understanding of the Nature of Science (NOS). This field aims both to trace how people think about science as well as to find ways of improving the accuracy and utility of those conceptions, particularly in science learners. The field of NOS research is characterized by debates regarding the constructs of NOS inquiry, the correct normative principles of science (against which to evaluate lay beliefs), and the best methods to elicit, study and teach these scientific norms. However, little research has inquired into the kinds of conceptions targeted by NOS inquiry itself, and few studies have surveyed more than a handful of NOS research tools. Similarly, little research has sought to uncover historical patterns of change in the conceptions underlying NOS constructs. This study therefore surveys the conceptions addressed over the history of NOS research, developing a comprehensive analytic framework to trace the scope of the topics and issues that feature in a diverse range of 81 NOS assessments, covering over six decades of instrumentation. By tracing historical change in the kinds of conceptions that have been targeted for investigation by NOS researchers, the study explores how debate about measures and norms has led to new and modified instrumentation. Finally, the analysis reveals targets of inquiry that have been insufficiently investigated, and which present productive new lines of inquiry.

Keywords: nature of science, epistemic cognition, personal epistemology

Introduction

Since at least the 1950s, the flourishing field of Nature of Science (NOS) research has investigated peoples' conceptions about science, or their scientific epistemologies. These conceptions are important because they influence the degree to which the public understand and accept the findings of scientific inquiry, as well as the ways in which people approach the learning of science. Over six decades, NOS researchers have sought to uncover what we believe about what science is, how science works, what scientists are like, and how scientific knowledge and practices are generated, evaluated and justified.

The field of NOS research features extensive debate about what constructs researchers should use for NOS inquiry, as well as the norms of science that should guide NOS investigation. Much of this debate has concerned the differing epistemological perspectives that underlie the design of rival NOS assessments. In spite of this, little research has systematically investigated the conceptions that are implicit in the history of NOS research. Similarly, little research has sought to uncover patterns of change in the conceptions underlying NOS constructs. This study aims to contribute to this line of inquiry, by systematically exploring the kinds of conceptions implicit in NOS instrumentation over the history of NOS inquiry.

The expanded framework developed in Chinn et al. (2011) defends a revised approach to modeling and measuring epistemic cognition, and this has implications for the study of scientific epistemologies. In developing this framework, Chinn et al. argued that a more fine-grained and comprehensive profile of a wider array of epistemic cognitions promises to improve the conceptual coherence and predictive validity of the field (Buckland & Chinn, 2015; Chinn et al., 2011; Chinn, Rinehart & Buckland, 2014). In the current study, the expanded framework provides an initial scaffold to explore the range of beliefs and attitudes about science targeted in NOS research. One important aim is to uncover changes in the focus of NOS research over historical time, as debate about measures has led to new and modified instrumentation. The Chinn et al. framework also reveals aspects of cognition about science that have been the subject of relatively little investigation, and thus suggests productive new lines of inquiry for the field.

The key research questions of this study are:

1. What are the range of topics and issues targeted by NOS assessments?
2. What is the distribution of assessment items across different NOS conceptions, both within and across studies?
3. How have the conceptions targeted for study changed over the history of the field?
4. Which NOS conceptions remain under-researched?

Research into the history of the field of NOS inquiry is important for a number of reasons. First, the field is sizeable, enduring and diverse, and a systematic overview is likely to help refine conceptions of the proper target of NOS inquiry. This is particularly important given the continued debate about what account of the norms of science should guide NOS inquiry. Second, multiple sources influence NOS conceptions, including psychological and educational research as well as the philosophy of science. Historical changes in NOS research conceptions may be the product of shifts in philosophy or psychology, but they can also be a product of the fashionable trends of an era. Changes in researchers' conceptions of the scientific enterprise are highly likely to impact on the ways in which they design NOS assessments. Tracing the conceptions implicit in NOS research can therefore serve to reveal the epistemological commitments regarding science that served to guide the history of NOS inquiry. Third, as new research targets replace the conceptions targeted by older instruments, it is valuable to identify inadequately investigated but important constructs. In developing a high-level and large-scale analysis of NOS assessments, one can more easily identify and investigate these under-researched NOS conceptions. It is therefore important to understand the historical trajectory of change in NOS-related conceptions, both for identifying the epistemological norms of science that underlie NOS research and teaching, and for refining and improving the constructs of NOS inquiry.

Scientific Epistemologies

Research into the epistemic cognition of science investigates 'folk', lay, novice, learner and expert conceptions of the character, norms and practices of science, i.e. the 'nature of science'. The field of NOS research aims to trace both inter-personal and intra-personal variation in 'scientific epistemologies', as well as to find out how these beliefs and attitudes are implicated in scientific understanding, acceptance and expertise. Some working in this field seek to develop accounts of the norms of scientific inquiry, for purposes of guiding science instruction and assessment. These researchers aim to specify what the nature of science actually is. For example, a 2003 study by Osbourne, Collins, Ratcliffe, Millar and Duschl's aimed to "determine the characteristics of scientific enquiry and those aspects of the nature of scientific knowledge that should form an essential component of school science curricula" (p. 1). Others, like McComas and Olson (1998), have instead tried to determine the norms and conceptions embedded within existing science education standards documents. A related strand of research has sought to trace the norms and conceptions of science that are implicit in the cognition of learners, teachers, experts and the public at large (e.g. Lederman, 2007). These researchers have targeted a wide variety of science-related topics, issues and concepts, including students' and teachers' understanding of laws, theories, experiments, methods and the goals of science, scientific change, progress and certainty, and socio-cultural influences on science and scientists. They have also conducted their research under a range of different terms. For example, Abd-El-Khalick and Lederman (2000) have targeted "conceptions of the nature of science," Nott and Wellington (1996) have investigated "views of the nature of science," and Sandoval (2005) has investigated "scientific epistemologies." However, they share the common goal of determining how people think about science, a question relevant both to science education and to the public understanding and support of science (e.g. Barufaldi, Bethel & Lamb, 1977; Kelly & Duschl, 2002; Osborne, 2002; Sadler, Chambers & Zeidler, 2004; Tao, 2003; Tobin & McRobbie, 1997; Waters-Adams, 2006).

NOS research instruments typically consist of written or interview assessments, generally comprising sets of statements and/or questions designed to elicit important cognitions about science and scientific knowledge (e.g. Abd-El-Khalick & Akerson, 2004; Bell, & Schwartz, 2002; Lederman, Abd-El-Khalick, Lederman & O'Malley, 1990). Some assessments investigate how participants respond to theoretical disputes in science, typically by presenting them with vignettes involving disagreement between scientific theories or experts (e.g. Lederman et al., 2002; Smith & Wenk, 2006). Others feature rich descriptions of authentic scientific contexts, and sometimes demand authentic inquiry from participants (such as the Conceptions of Scientific Theories Test by Cotham and Smith, 1981). Assessments also vary in terms of the degree to which they assume a specific disciplinary or pedagogical context, as well as the scope of the conceptions about science that they are designed to trace. While some aim more narrowly (e.g. only on conceptions about scientific theories in high school Physics education), others target a much broader range of science concepts. A subset aims to trace conceptions that are of particular importance for understanding how science learners achieve a normative conception of science (e.g. Khishfe & Abd-El-Khalick, 2002). These contrast with instruments that explore a broader range of conceptions, many of them non-normative and not correlated with success in science understanding. Although the former project is of particular educational import, both are valuable for the field. The latter is vital in that it includes a focus on conceptions that, while not necessarily educationally relevant, might play an important role in general reasoning about science.

The current study meta-analytically investigates the conceptions inherent in the NOS field. While several studies have conducted similar kinds of investigations – for instance, Lederman et al. (1998) surveyed 25 assessments, while Guerra-Ramos (2012) surveyed 19 – the current study is distinctive in including a far wider range of studies than has previously been surveyed. Abd-El-Khalik (2013) analyzed 32 peer-reviewed instruments, developing a score of their empirical use, reuse and transmission amongst researchers. This analysis was primarily methodological in nature, assessing the patterns of change in measurement methods over the history of the field. It also compared instruments in terms of their elicitation of different clusters of science concepts (e.g. science as tentative, empirical, etc.). However, it did not describe the distributions of these clusters across the instruments studied. In contrast, the current study aims to trace the conceptual (rather than methodological) structure of NOS instrumentation, and tends towards inclusivity of instruments. Even slightly modified versions of existing NOS instruments were thus included in this survey in order to maximize its completeness.

Much NOS research characterizes the conceptualizations of science learners, teachers and in the public at large in terms of a deficit model. For example, the “Views of Nature of Science” questionnaire in Abd-El-Khalick (2001) revealed student teachers as failing to recognize the essentially inferential, tentative, theory-laden and creative nature of science. In response to this trend, considerable debate in the field has focused on the degree to which these deficit findings are accurate, pedagogically significant, or relevant to actual inquiry (e.g. Allchin, 2011; Nott & Wellington, 1998; Sandoval, 2005). A closely related methodological disagreement among NOS researchers concerns the structure that cognitions about science exhibit, and consequently the way in which they are best elicited and studied. On the one hand, many studies appear to reveal participants' scientific epistemologies as involving relatively stable, coherent systems of explicit and general beliefs. On the deficit model, these beliefs tend to be naïve, shallow, over-simplified and inaccurate. Nonetheless, researchers have assumed that they are reflectively accessible to learners and thus elicited relatively easily by assessment items (e.g. Rampal, 1992). Others have developed accounts on which people demonstrate a fragmentary, inconsistent, conflicted and even radically unstable scientific epistemology (Elby & Hammer, 2001; Hammer, 1994a). For these researchers, the observed coherence of participants' epistemic conceptions is a product of the design of instrumentation, rather than a reflection of the structure of their beliefs. The call for investigation into the consistency, fragmentation and stability of real epistemic cognition is important and long overdue. However, the way in which researchers carve up the epistemic landscape is likely to impact on their attributions of coherence or fragmentation. Studies like the current one, which aim to trace these systems of individuation, are therefore vital.

While some NOS studies assume a set of normative foundations for what counts as a sophisticated and accurate science conception, others are primarily descriptive, merely charting the range of participant conceptions. For example, Driver, Leach, Millar and Scott (1996), developed an analytic framework inductively by identifying coherences across participants' science discourse, without presupposing criteria for science conceptions. In contrast, Kimball's (1967) early "Nature of Science Scale" focused on eight core normative NOS principles, developed from the views of experts. However, considerable debate has also centered on what constitutes a better or more accurate account of the true nature of science as well as the right account of the norms and misconceptions of science (e.g. McComas, Clough, & Almazroa, 2000; Leach, Millar, Ryder & Séré, 2000). Researchers have proposed varying sets of fundamental tenets of good science (e.g. Duschl, 1985, 1988; Hodson, 1988), some of which have then guided the design of NOS instrumentation (e.g. Lederman, 1985; McComas et al., 2000). However, given the enduring diversity of views amongst philosophers on this question (e.g. Boyd, Gasper & Trout, 1991), there seems to be little prospect for consensus. The current meta-analysis therefore avoids this debate by considering the full range of conceptions targeted in NOS research without specifying a normative account of science in advance.

In sum, as Abd-El-Khalik (2013, p. 40) puts it, "the seeming dominant narrative in the field is one that is focused on a plethora of NOS assessments and continued disagreement about the construct." This study aims to contribute to the resolution of some of this disagreement by providing a comprehensive account of the conceptions that feature across the history of the field of NOS research. It also aims to identify important NOS cognitions that have thus far been under-researched, and to develop tools with which to construct and evaluate new instrumentation.

Method

Selection of NOS Surveys

The 81 assessments surveyed in this study featured in publications that explicitly aimed at charting NOS conceptions. Studies that used modified versions of existing assessments were included, regardless of the degree of modification. For example, the Schwartz (2004) study combined components of the VNOS-C and VNOS-B with other items, generating the VNOS-Sci instrument; all three of these surveys were thus included. Most studies incorporate at least some items from previous studies, either in a duplicated or modified form. The only precondition for inclusion was that the candidate NOS instrument differs from an existing survey.

A number of library and electronic searches served to identify candidates for inclusion. For example, the Google Scholar search tool identified the following combination of key words: "nature of science," "survey," and "assessment." This search yielded 9800 hits in February 2012. Examination of these results then determined whether they referred to any unique NOS assessments. In addition, the Google Scholar search tool conducted an automated weekly search for newly published NOS research for two years. Researchers searched all identified articles for reference to NOS surveys, with this procedure continuing iteratively to uncover further sets of instruments. Special attention was paid to meta-analytic surveys (e.g. Abd-El-Khalik, 2013; Guerra-Ramos, 2012; Lederman, 2006, 2007), which describe and review numerous assessments and therefore represent a rich resource for identifying NOS instrumentation.

The sample excluded instruments for a variety of reasons. First, the mere replication of the items of a previously published study served as a basis for exclusion. Second, failing to include any explicitly conceptual or linguistic items (e.g. the “Draw a Scientist” studies of Chambers, 1983; Schibeci & Sorenson, 1983) also served to exclude assessments. Third, focusing specifically on practices or attitudes, rather than articulable beliefs about science, provided additional grounds (e.g. Nott & Wellington, 1998; Swan, 1966). Finally, engaging participants in active inquiry, rather than eliciting beliefs about science, excluded assessments (e.g. Windschitl, 2004). Some popular NOS instruments were not included because they were unpublished, out of print, or promulgated by since-defunct institutions (e.g. the ROSE studies).

Table 1 presents the full list of assessments analyzed, along with the number of items comprising each. Surveys were included even if researchers could only recover a partial list of its constituent items. The study therefore aims to capture NOS instrumentation over the history of the field in as comprehensive manner as possible.

Inter-Rater Reliability

A main coder who was blind to the name, date or developers of each survey conducted several rounds of initial coding. During this process, the coding scheme was developed and refined. Two assistant coders then practiced assigning codes to a randomly selected set of practice surveys. Once familiar with the scheme and proficient at assigning codes, the assistants coded a randomly selected and blinded subset of the remaining surveys that ranged between 17% and 24% of the full data corpus. The arithmetic level of accord between main and assistant codes assigned to the selected subset of surveys served as a measure of inter-coder reliability.

The first assistant coded the following categories: *Aims and values*, *Epistemic virtues, vices and responsibilities*, *Non-NOS* and *Follow-up*, attaining an agreement of 96% for the set.

The second assistant coded the *Nature and Structure of Scientific Knowledge* category, obtaining 94% agreement for this set, and subsequently assigned the codes from the remaining categories, obtaining 91% agreement for this set. Overall, the high level of agreement between coders on the selected surveys shows the consistent application of the analytic framework.

Coding Strategy and Framework

Once the assessments were gathered, the resulting analysis involved two main phases. First, the expanded framework explicated in Chinn et al. (2011) provided the initial structure for categorizing the conceptions addressed by the NOS instruments. Second, a descriptive synthesis of the subcomponents of the framework inductively identified patterns exhibited by multiple items. Thus, while the Chinn et al. framework provided a useful initial tool for high-level categorizing, the sub-components were formulated in a primarily bottom-up and inductive manner.

In developing the analytic framework, examination of each assessment item served to reveal its central target cognitions. For example, questions about the methods of science differed from questions about its conceptual basis, as well as from questions about its social and institutional structures. This process resulted in the seven overarching categories of the framework: *Scientific aims and values*; *Nature and structure of scientific knowledge*; *Sources, justification and epistemic stances of scientific knowledge*; *Epistemic virtues, vices and responsibilities of scientists*; *Reliable processes for achieving epistemic aims*; *Non-Nature of Science*; and *Follow-up*. Items falling into these overarching categories received one or more codes, depending on the range of conceptions that they featured. Table 2 provides a more detailed overview of the coding scheme, showing the seven overall categories, the eighteen major subcategories, and the 113 individual codes. Tables 3 to 8 present each of the codes and overarching code categories in more detail.

Scientific aims and values. Table 3 presents the category capturing items that focus on the aims and values of science, including both epistemic and non-epistemic aims. The category includes the valuations that serve to ground claims about scientific aims, i.e. references to what makes science valuable. The subcomponents comprise both generic specifications of aims, as well as a range of the more specific epistemic aims, including truth and the revealing of reality, the discovery of new phenomena, knowledge, explanation and understanding as goals, and the formulation of laws and theories through the systematization of observation. This category also identified various non-epistemic aims of science, including the improvement of human welfare, the control of nature, practical benefits like technological inventions, the honors, grants and rewards that motivate scientists, and scientists' religious or spiritual aims. The final subcomponent captured the non-epistemic, ethical value of science, including its moral evaluation, the ethical and professional systems within science, and the sanctity of nature.

Nature and structure of scientific knowledge. Table 4 presents the category concerned with defining and characterizing the nature of science, specifically capturing the kinds and distinctive characteristics of scientific conceptions. The *Defining and demarcating science* sub-category captures generic definitions or requests for definitions of science, as well as broad descriptions and analogies involving science. This sub-category also addresses the extent and limitations of science and relations between science, religion and the supernatural. The *Kinds of scientific conceptions* sub-category captures items focusing on the variety of conceptual structures that science features. The *Characteristics of scientific conceptions* sub-category identifies descriptions of these conceptions. Finally, *Change in scientific conceptions* identifies items oriented on scientific change, capturing generic references to the mere possibility of change as well as descriptions of specific kinds and causes of change.

Sources, justification and epistemic stances. Table 5 identifies items focused on the origins and justification of scientific knowledge, including the epistemic stances adopted towards it (e.g. certainty, doubt, provisional acceptance, etc.). This category captures items targeting both generic and specific sources of scientific knowledge. It also captures items oriented on disagreement in science, including the possibility and significance of rival conceptions, the reasons and sources of scientific disagreement and the achievement of consensus. The *Scientific justification* sub-category captures items oriented on generic as well as topic-specific justifications of science. This includes issues involving the legitimation of scientific claims, the status of scientific anomaly, and the “theory-laden” nature of observation. The *Epistemic stances* sub-category captures the reflective attitudes that scientists might adopt with regards scientific knowledge, including attitudes of certainty, uncertainty, tentativeness and conjecture.

Epistemic virtues, vices and responsibilities. Table 6 captures items that describe the typical, important or distinctive characteristics, traits, competences, abilities or dispositions of scientists (in both positive and negative forms), as well as those describing science itself in an explicitly cognitive way. The first sub-category captures the distinctively epistemic features of scientists, i.e. the characteristics that relate to their role as knowledge-producers. This category did not capture references to the faculties of scientists (e.g. their perceptual capabilities), although claims that specific traits, beliefs or abilities are necessary for scientific competence were included. It also captures items that address scientists’ assumptions regarding the intelligibility and order of the world they investigate, as well as the styles of cognition that typify science. The second sub-category, *Other characteristics of scientists,* captures any non-epistemic descriptors of scientists, i.e. (e.g. their race, religion, etc.) or ambiguous terms (e.g. “eggheads”). This includes descriptions of scientists as religious, non-religious or anti-religious, as well as their philosophical, ontological and epistemological commitments.

Reliable processes for achieving epistemic aims. Table 7 captures all references to the methods and processes of science, as well as its social and institutional dimensions. This category includes items that target participants’ beliefs about the many processes implicated in the production of scientific knowledge. The *Characteristics of scientific processes and methods* sub-category captures generic references to the processes and methods of science. It also includes the role, utility and value of inaccurate conceptions for scientific methods and the degree to which scientists are free or constrained in their research practices. The *Kinds of scientific processes and methods* sub-category captures items that invoke specific methods, while other codes include criteria for the evaluation and selection of conceptions, as well as the role of luck and serendipity in science. The *Social and institutional processes* sub-category captures items that refer to the scientific community, including the role of publication, citation, reputation and communications. Also captured here are generic references to the influence of social, economic, cultural, religious, ethnic, political and national values and factors, as well as the impact of legal, governmental, corporate and special interest groups. Examples include reference to gender roles and feminism, as well the social decision-making power of scientists.

Non-NOS and Follow-up. Table 8 captures items that do not directly concern the nature of science, but focus instead on the teaching or learning of science, or on affect and attitudes. This category also targets participant beliefs about the necessity, validity and practical usefulness of science learning, and includes items that target scientific literacy in which participants make judgments about specific scientific claims. While this class of judgments represents an important target of psychological research, they either involve attitudes rather than beliefs (e.g. the enjoyment of science) or focus on science learning and teaching rather than the nature of science itself. Also listed in Table 8 were *Follow-up* items, which required participants to explain, defend, exemplify, elaborate or justify their responses to prior questions. This was included as a distinct category because of repeated calls through the history of NOS research for more elaborated and participant-centric measures.

Results

The Surveys

Figure 1 shows the historical distribution of surveys, from 1950 to 2012, with the height of each data bar indicating the number of items within each assessment. While there were intermittent studies conducted between 1950 and 1995, the late 1990s to the present has seen an explosion in the number of studies targeting NOS conceptions. As around 75% of the surveys post-date 1996, the data corpus is heavily weighted towards research conducted in the last two decades.

Overall Distribution of Coding Categories

The 81 assessments surveyed presented participants with 2528 individual items, with an average number of 31.2 items per survey, ranging from a low of four items to a high of 135. The modal number of items was eight, and the median was twenty-three items. The total number of codes assigned to the items was 3372, at an average of 41.6 codes assigned per survey and 1.33 codes per item. Figure 2 shows the distribution of codes across the seven major code categories. The *Scientific aims and values* category constitutes the smallest category, while *Nature and structure of scientific knowledge* is the largest.

Figure 3 shows the breakdown of the codes assigned into the 18 sub-categories themselves, revealing that social processes of science constitutes the largest sub-component of the *Reliable processes for achieving epistemic aims* category. In addition, the *Epistemic virtues and vices* sub-category significantly outweigh the non-epistemic virtues and vices within the over-arching category. The *Non-NOS* category, which identified items excluded from consideration, captures almost a fifth of all items.

Tracing historical change in NOS surveys

Tables 3 to 8 show the distribution across the full corpus of data. Each code is associated with three data points. Column A presents the percentage of the total number of items assigned that code. Column B presents the quantity of items assigned that code as a percent of the coding category in which the code falls. Column C presents the percentage of assessments themselves that received each code.

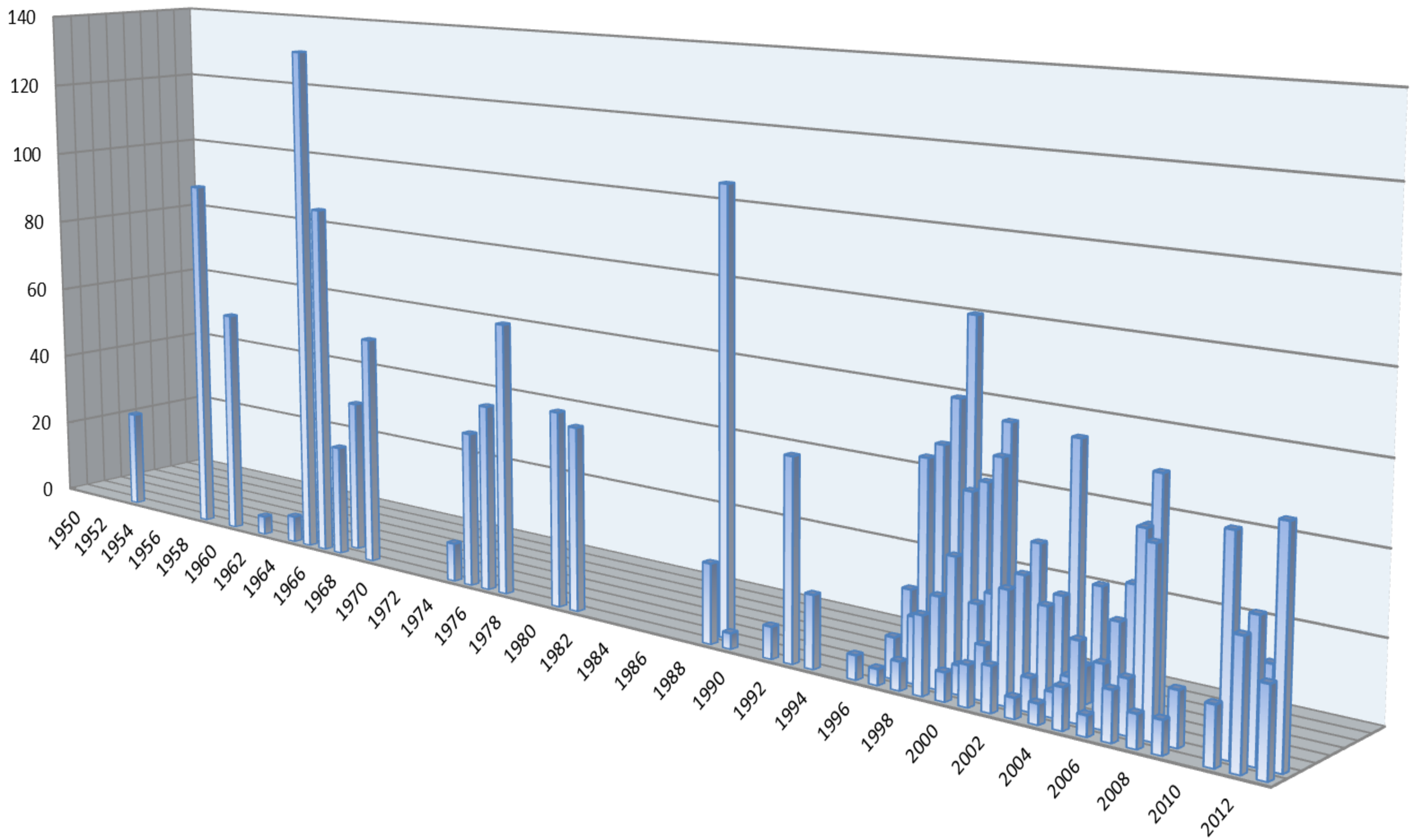


Figure 1. Historical distribution of NOS instruments. This figure shows the targeted NOS surveys by year, with the height of the data bars indicating the number of items included in each survey, from 1950 to the present.

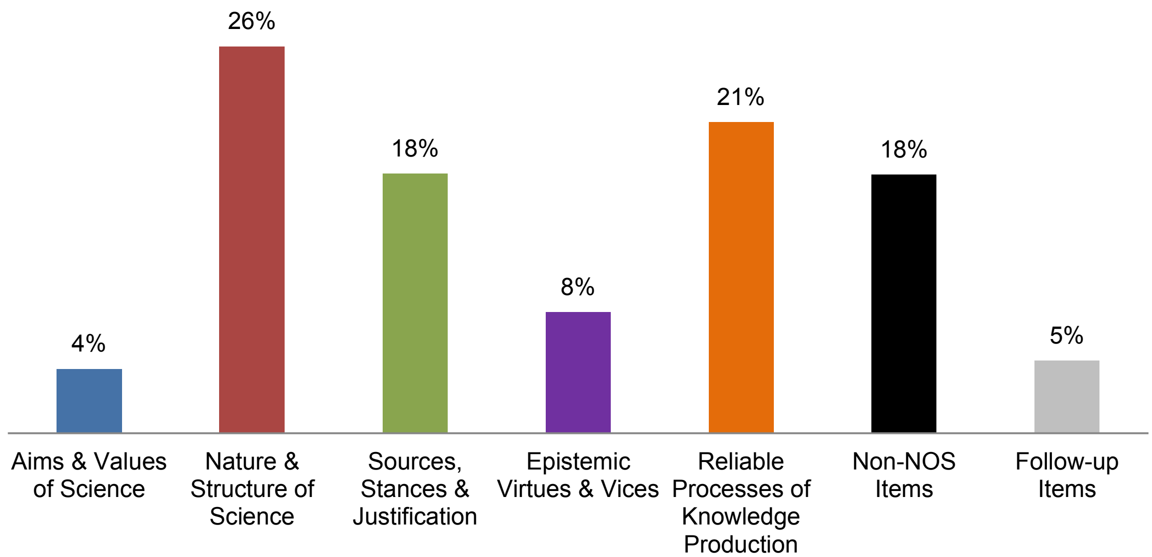


Figure 2. Coding distribution for the seven code categories. This figure shows the total percentage distribution of codes across the seven overarching code categories.

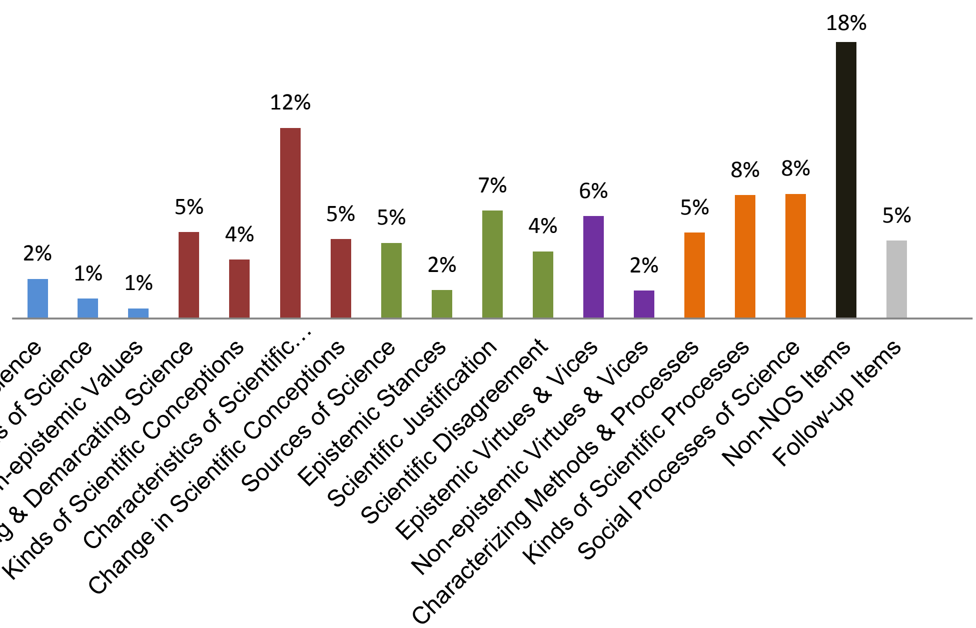


Figure 3. Coding distribution for 18 sub-categories. This figure shows the total distribution of codes across the 18 sub-categories, with the major categories grouped.

Figures 4 to 13 capture the absolute or proportional relationships between each of the major categories and several of the sub-categories of the coding scheme, with every code included. The data reveals some interesting historical patterns of change. Figures 4 and 5 demonstrate that the *Sources, justification and epistemic stances* category had a very limited start in early NOS research in the 1950s and 1960s. By the 1970s, however, this category had expanded dramatically, with a large infusion of items specifically targeting issues around scientific disagreement and justification. By the 2000s and 2010s, this category had come to comprise a sizable proportion of all NOS items developed, with representation across all four sub-categories. The data also shows that while the *Epistemic virtues, vices and responsibilities of scientists* category begins as a substantial portion of the items featuring in early research, at least half focused directly on non-epistemic virtues and vices. The next two decades reveal a similarly substantive, though declining, focus on non-epistemic virtues and vices. By the 2000s and 2010s, assessment items targeting this area of NOS conceptions had shifted to focus almost exclusively on distinctively epistemic virtues and vices.

Figure 4, which includes absolute values for the major categories over time, shows the great profusion of research developed after approximately 1996. It also reveals that the *Nature and structure of scientific knowledge* category was routinely the largest across all decades, whereas the relationship between the other categories changed considerably between decades. For example, the *Sources, justification and epistemic stances* category was significantly larger than the *Epistemic virtues, vices and responsibilities* category in all decades except the 1950s. Figure 5 reveals how the proportionate distribution of codes over the sub-categories has changed dramatically over the six decades of NOS research.

Figure 6 shows changes within the *Scientific aims and values* category itself. This figure shows that early NOS research featured a substantial focus on distinctly non-epistemic aims. In particular, this focus was on the practical and technological applications of science, as well as on its religious, spiritual and moral value. In later decades, a far greater variety of aims came to characterize the category. In particular, distinctively epistemic types of aims came to dominate, with considerably less attention devoted to the ethical aims and values of science.

Figures 7 to 9 present the sub-categories that fall in the *Nature and structure of scientific knowledge* category. Figure 7 shows the total proportion of codes assigned to the *Defining and demarcating science* sub-category. It reveals that while the request for generic definitions of science have featured across all decades of NOS research, questions oriented on the epistemic status of science have grown significantly. Few items targeted the relation between science and technology before the 1980s, with the early decades focused instead on the relations amongst scientific disciplines. Figure 8 shows the *Characteristics of scientific conceptions* sub-category, revealing that a focus on how science has served to help or harm has been an enduring focus of NOS research, though one that is largely absent in the last decade studied. Figure 9 examines the *Change in scientific conceptions* sub-category and reveals a dramatic tapering of items targeting issues of scientific progress, which formed a substantive part of early research. This shift accompanies the steadily growing attention directed at the reasons and causes for change.

In Figure 10, the *Sources, justification and stances* category shows an increasing variation and specialization in the kinds of questions posed over time. Items in this category concentrated largely on the nature and character of perception, and the proportion of items concerning rival versus cohering scientific conceptions increased steadily. Figure 11 shows the distribution across the *Epistemic virtues, vices and responsibilities of scientists* category. One dramatic trend has been an increasing focus on creativity, imagination and intuition as a trait of scientists. In the 2000s, this grew to 30% of the epistemic virtues and vices investigated. Non-epistemic descriptors of science dominate this category in early decades of NOS research, yet these references steady declined, disappearing by the 2000s.

Figure 13, on the *Social and institutional processes of science* sub-category, shows that items oriented on issues of scientific community consistently represent around 10% of items across the decades. However, during the 1990s, this proportion increased to fewer than 50%. In addition, while items on the co-influence of social, political and economic forces on science featured were a small but regular feature across the decades, attention on these issues increased dramatically after 2000. The 1970s was an atypical decade across many of the categories and showed an outsized focus on issues of science funding and public support of science.

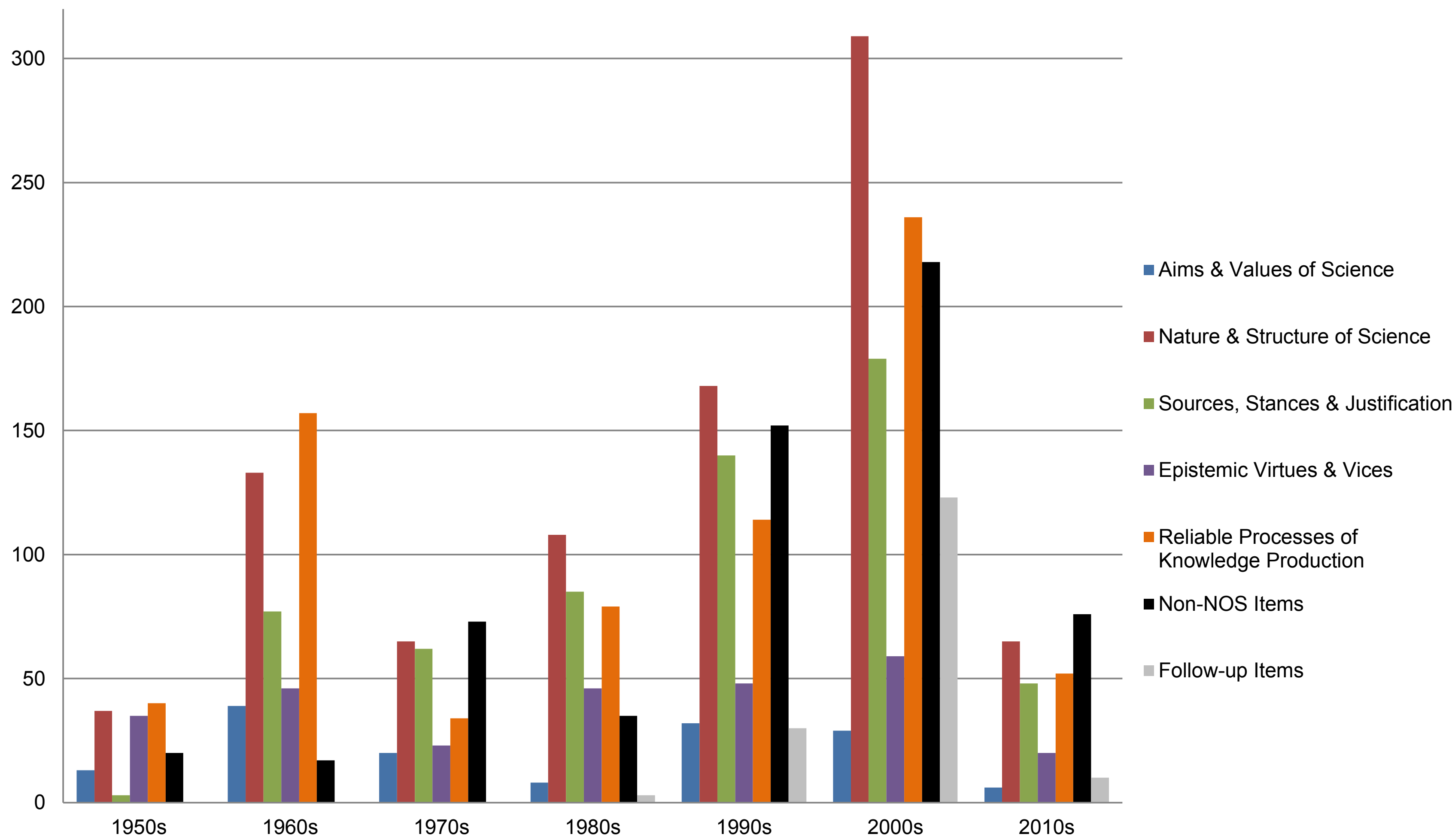


Figure 4. Distribution of categories by decade. This figure shows the total number of codes assigned to the seven major categories, over the six decades of NOS research.

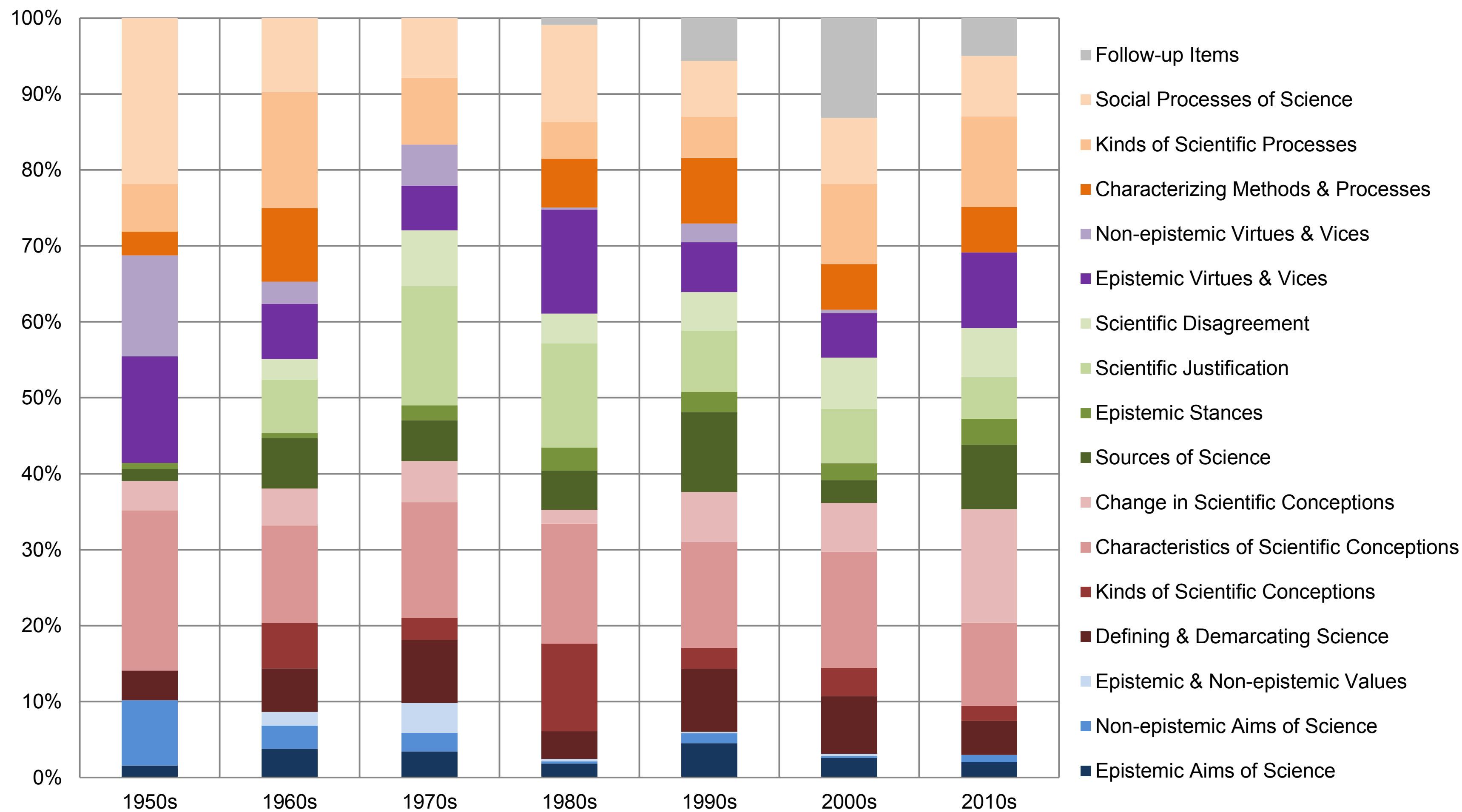


Figure 5. All NOS sub-categories by decade. This figure shows the total proportion of codes assigned to the seventeen sub-categories over the six decades of NOS research (excluding the Non-NOS category).

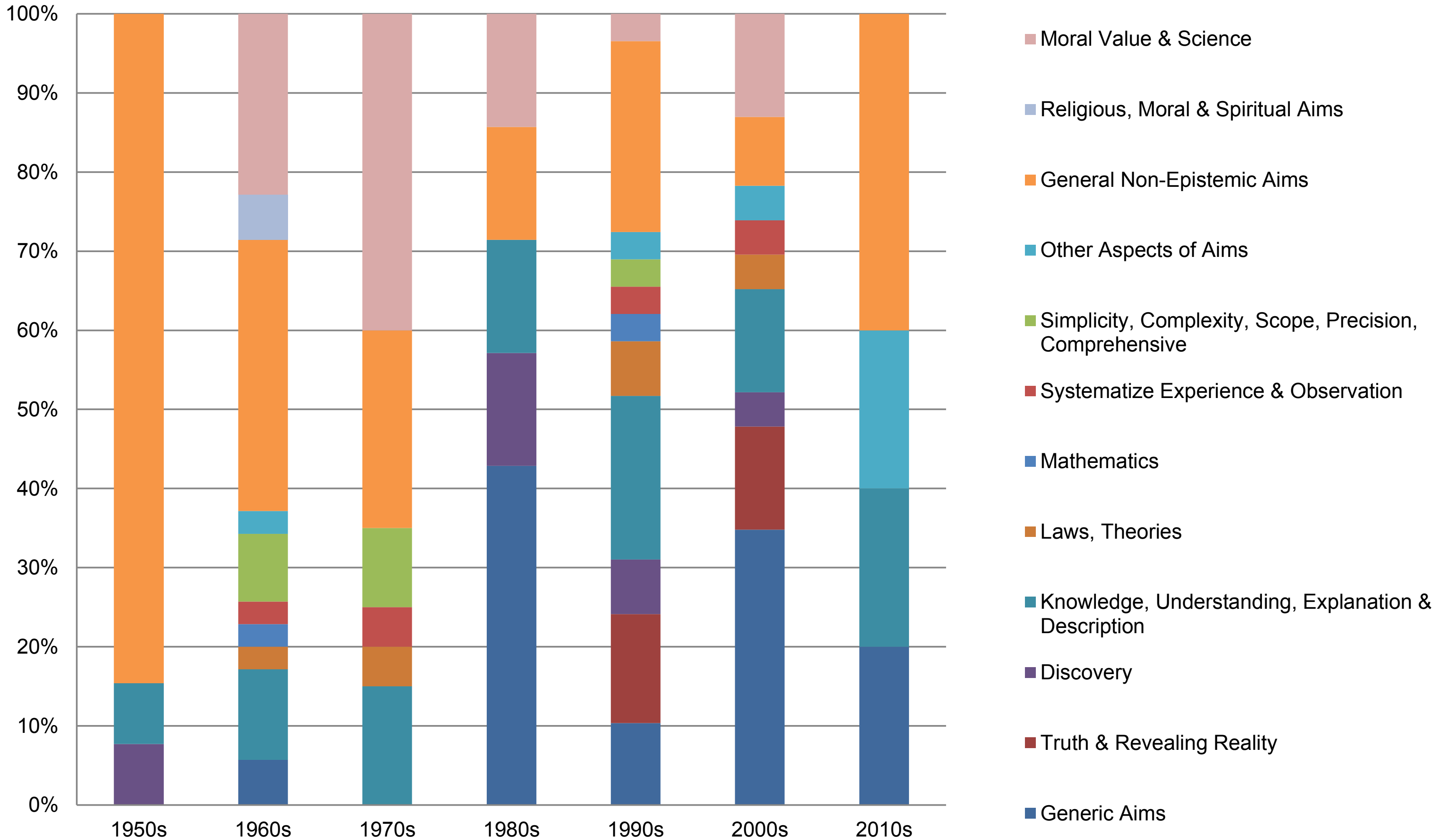


Figure 6. Scientific aims and values. This figure shows the total proportion of codes assigned to the *Scientific aims and values* category over the six decades of NOS research.

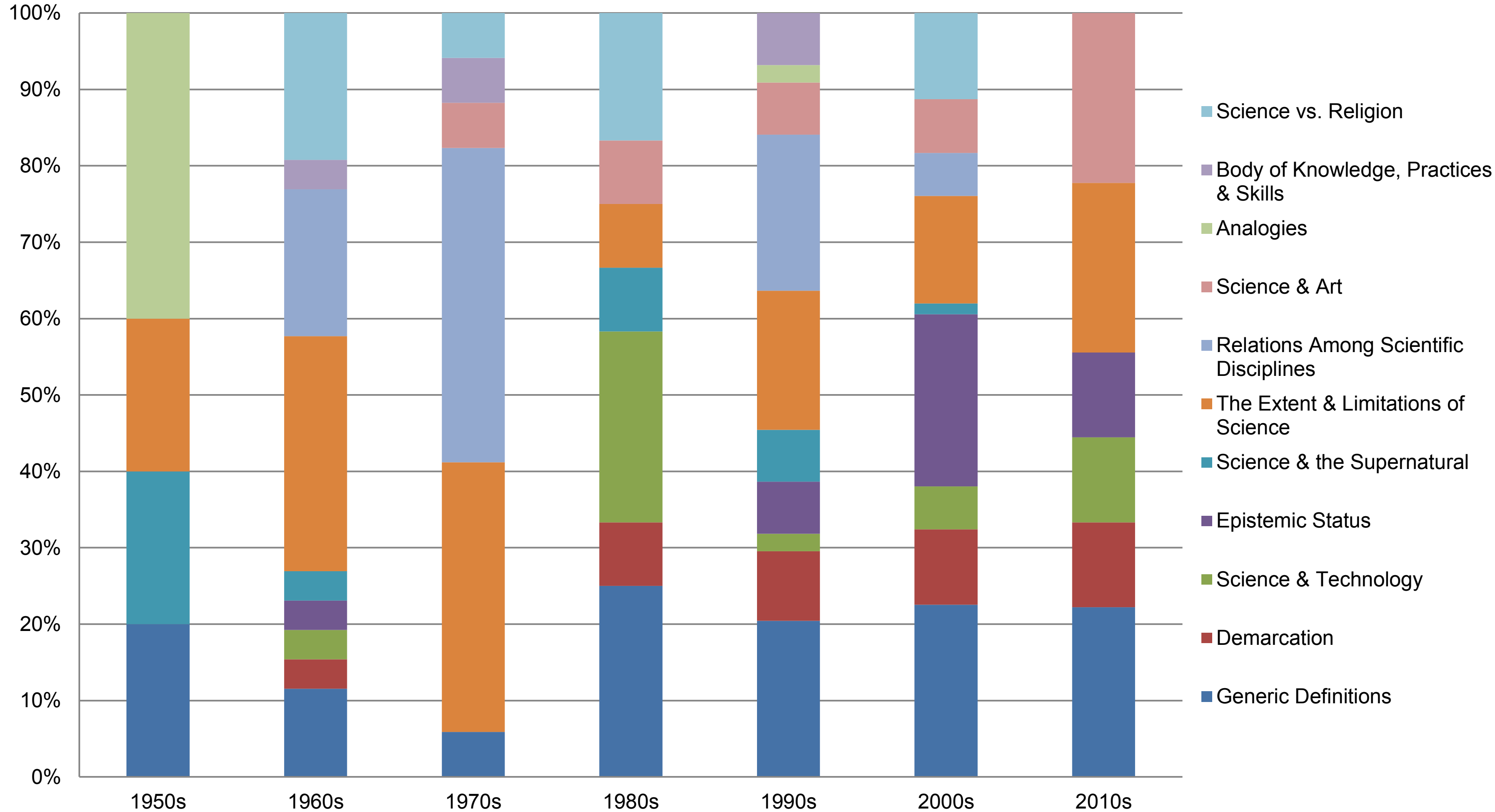


Figure 7. Defining and demarcating science. This figure shows the total proportion of codes assigned to the *Defining and Demarcating Science* sub-category over the six decades of NOS research.

Figure 8. Kinds of scientific conceptions and *Characteristics of scientific conceptions*. This figure shows the total proportion of codes assigned to these sub-categories over the six decades of NOS research.

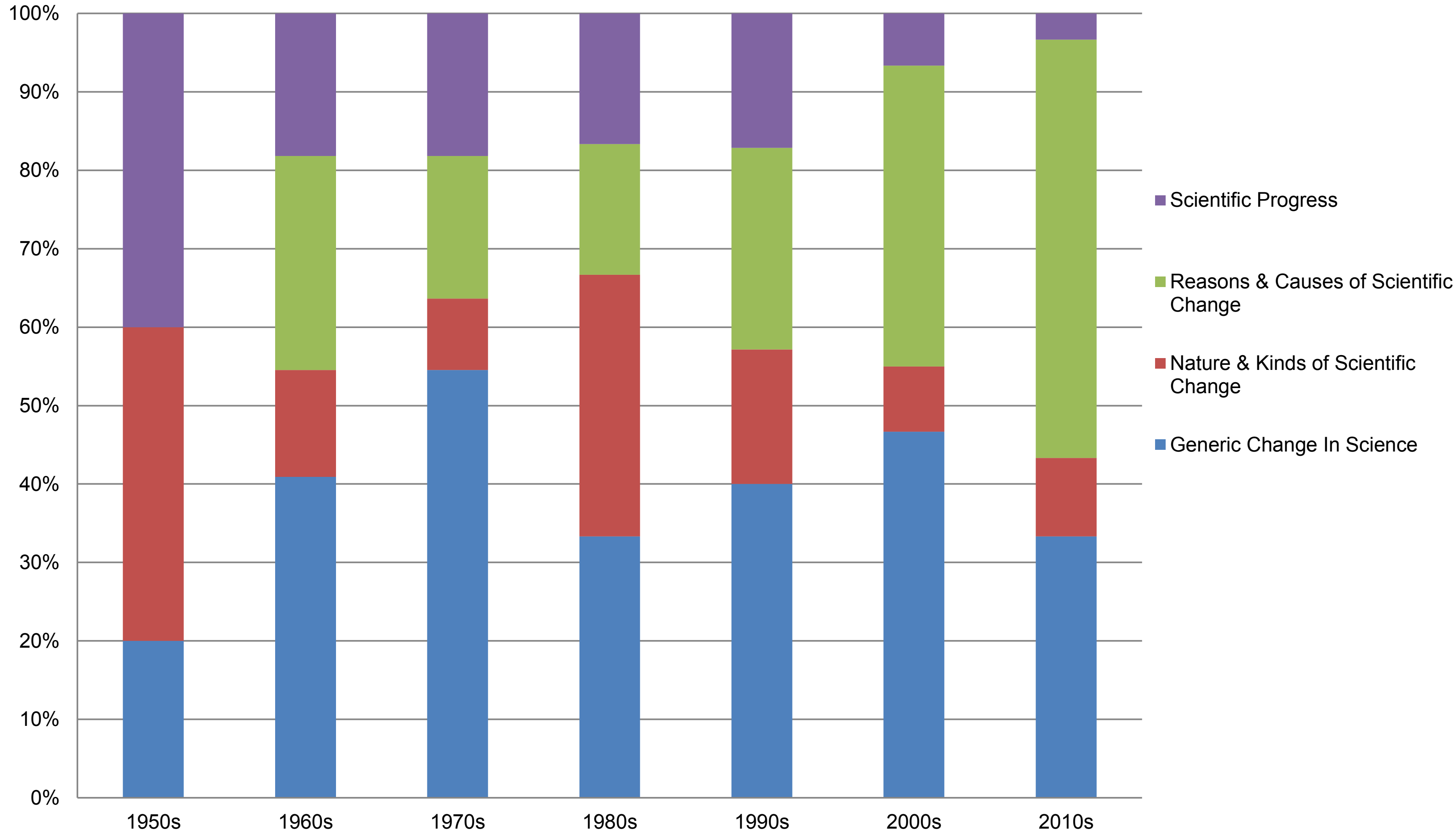


Figure 9. Change in scientific conceptions. This figure shows the total proportion of codes assigned to the *Change in scientific conceptions* sub-category over the six decades of NOS research.

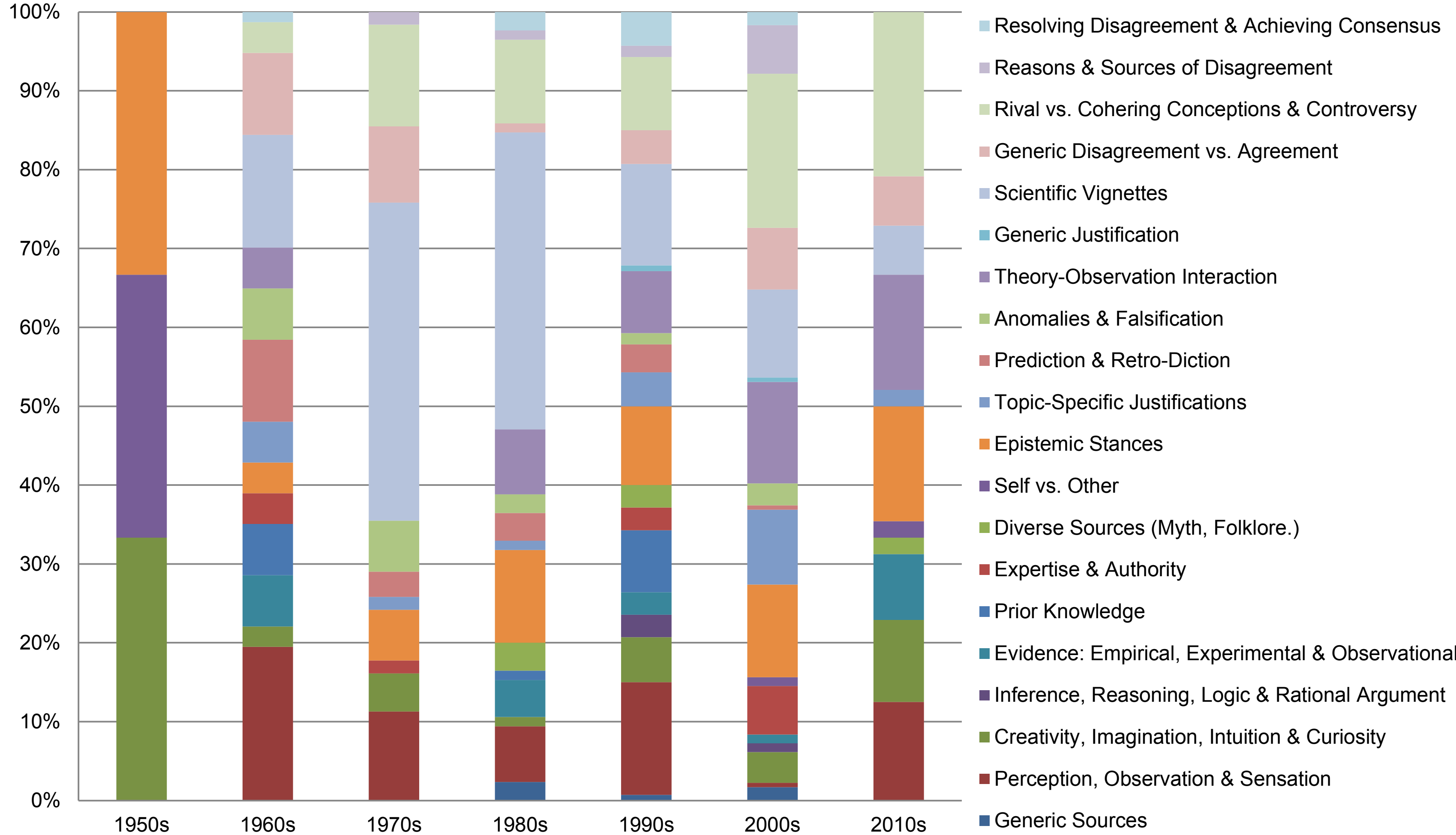


Figure 10. Sources, justification and epistemic stances. This figure shows the total proportion of codes assigned the *Sources, justification and epistemic stances* category over the six decades of NOS research.

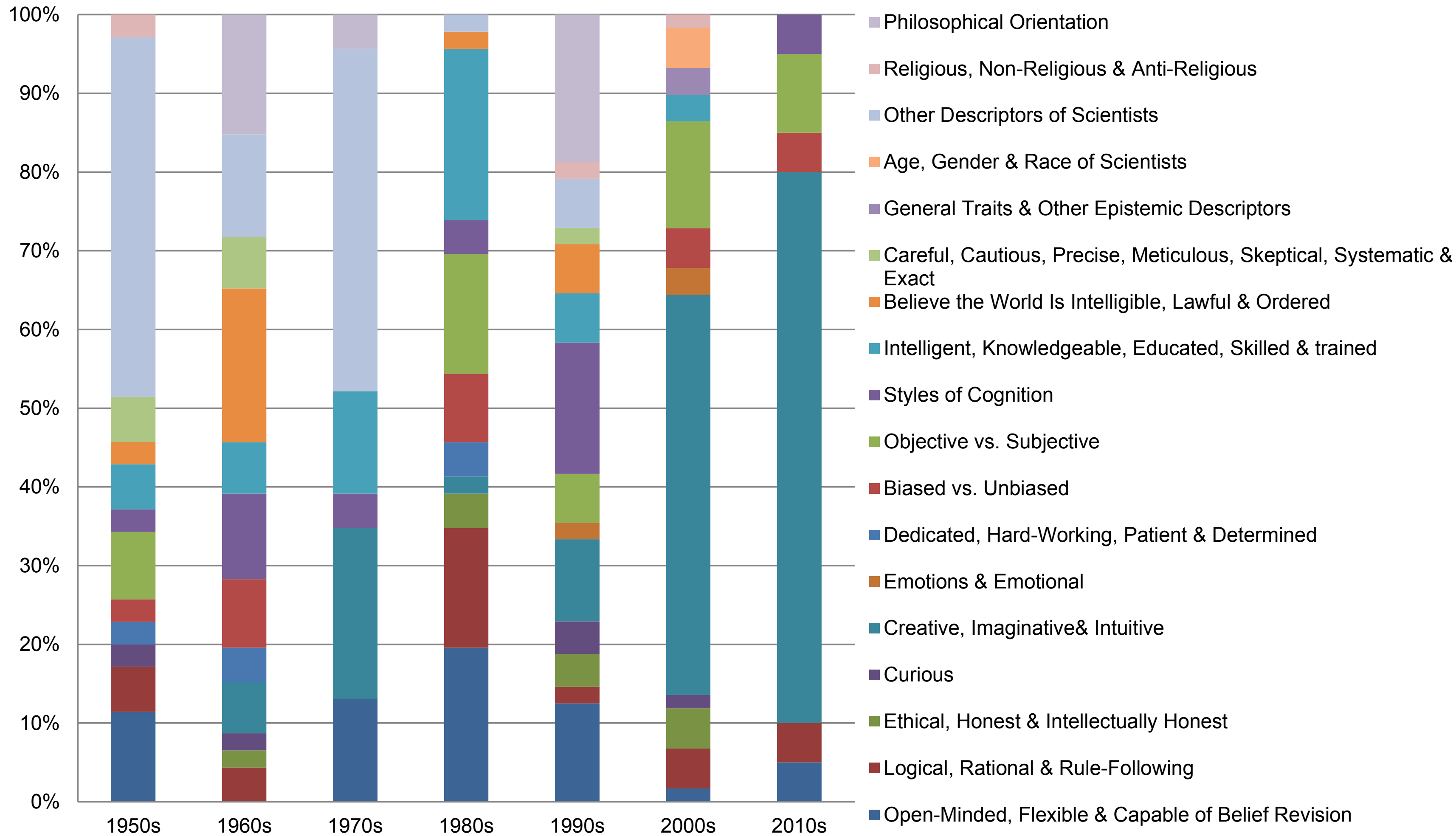


Figure 11. Epistemic virtues, vices and responsibilities of scientists. This figure shows the total proportion of codes assigned to the *Epistemic virtues, vices and responsibilities of scientists* category over the six decades of NOS research.

Figure 12. Characteristics… and Kinds of scientific processes and methods. This figure shows the codes assigned to two subcategories within the "Reliable processes of knowledge production" category.

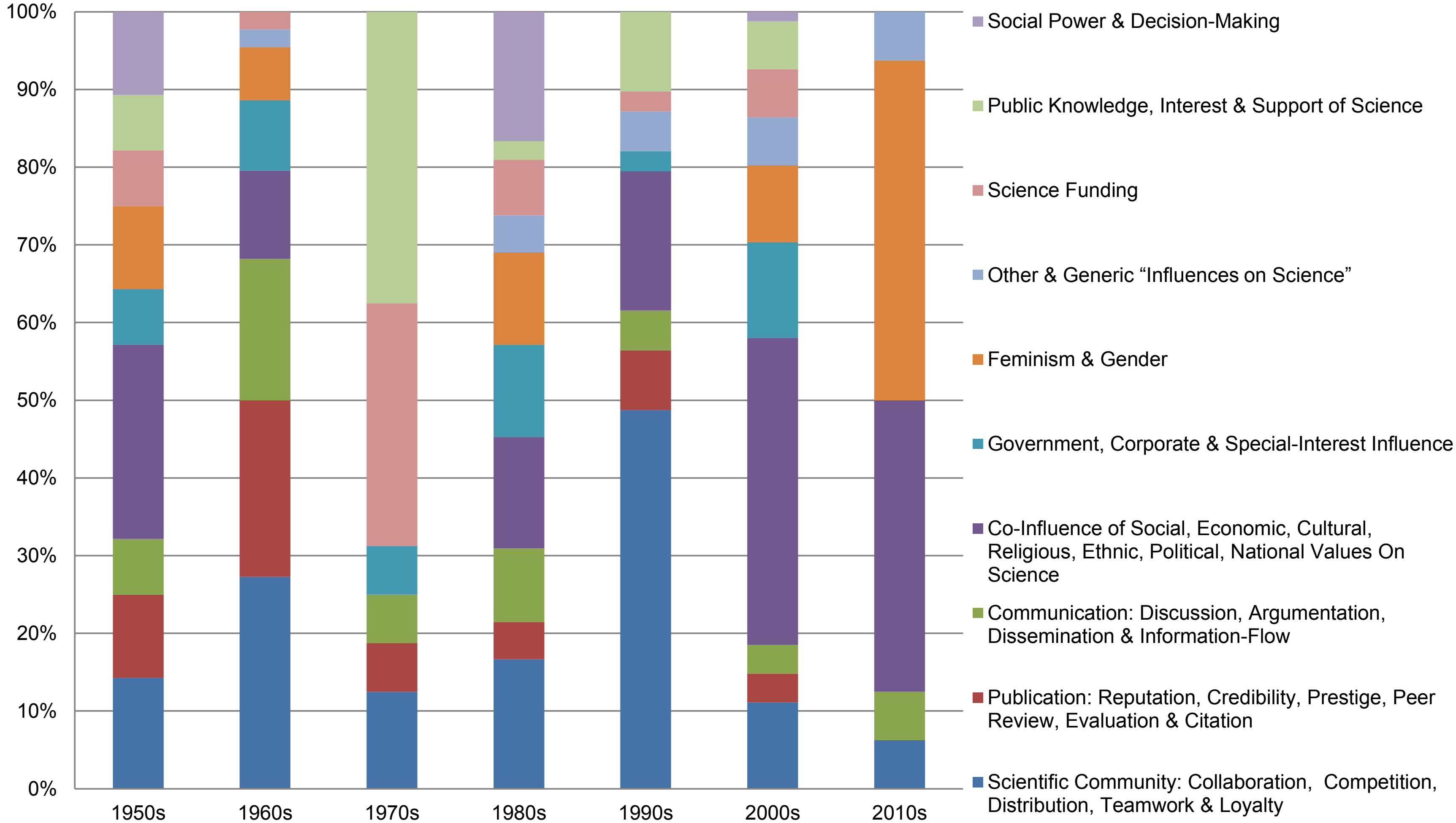


Figure 13. Social and institutional processes of science. This figure shows the total proportion of codes assigned to the final subcategory within the *Reliable processes of knowledge production* category.

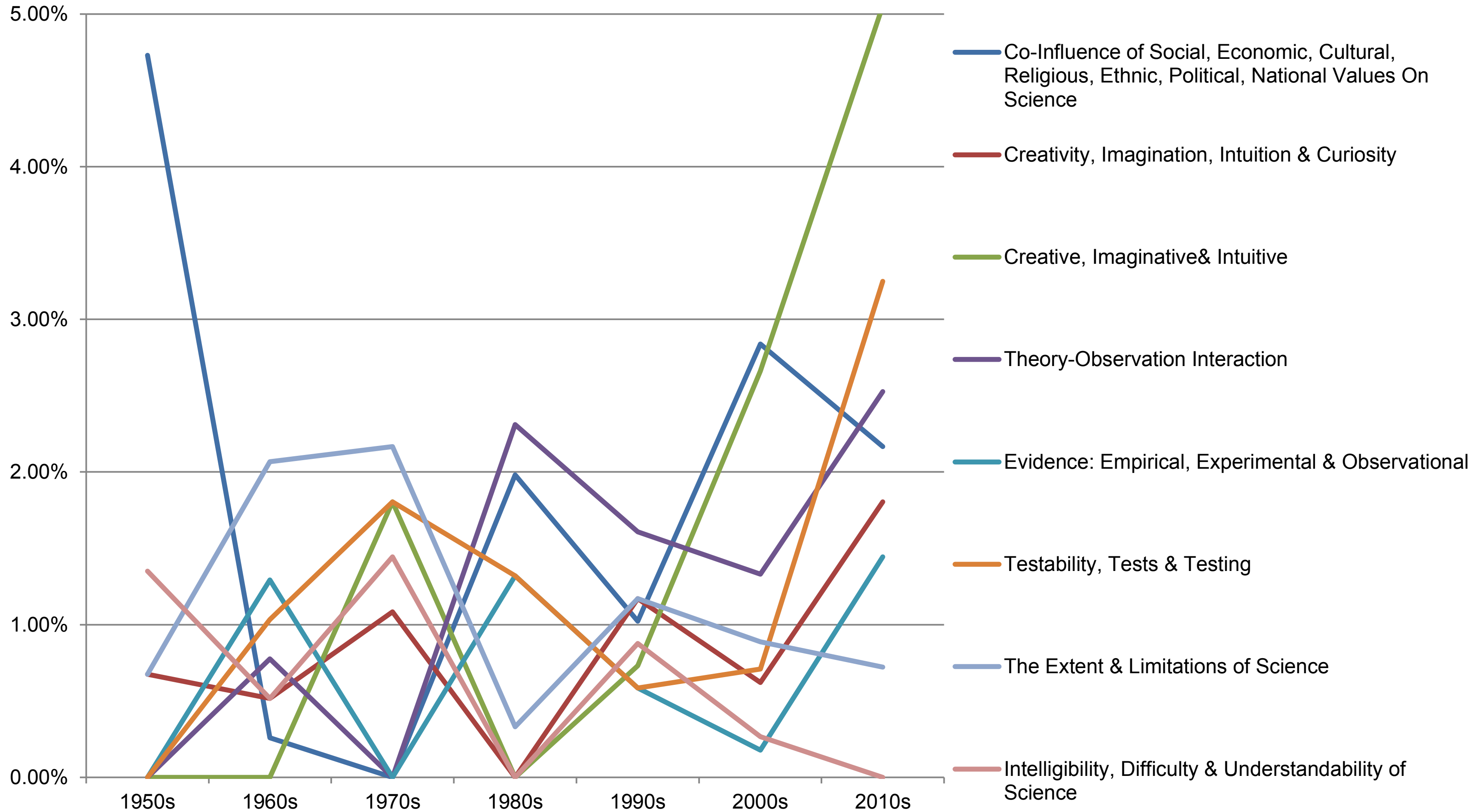


Figure 14. Change in distribution of selected codes across the surveyed period. This figure shows the proportion of codes assigned across all NOS assessments over the surveyed period, for selected code categories.

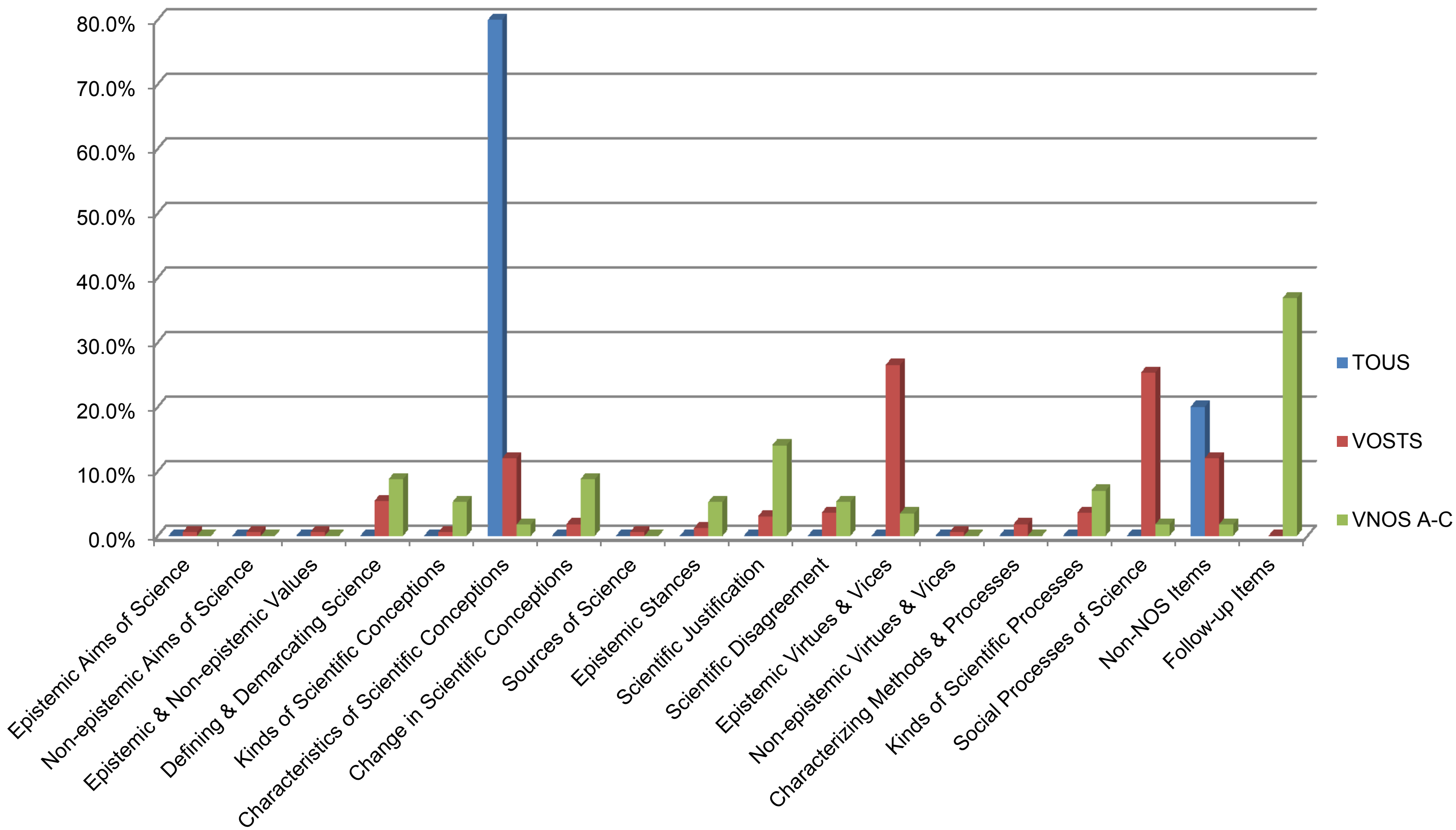


Figure 15. Comparative distribution of three NOS instruments across the sub-categories. This figure shows the proportion of codes assigned the subcategories for three selected NOS assessments: TOUS, VOSTS and VNOS A through C.

Finally, Figure 15 shows a snapshot of the difference in NOS conceptions targeted by three instruments, the TOUS, VOSTS and the combined VNOS versions A through C, across all the subcategories (see Table 1). Selection of these followed from Abd-El-Khalik's calculation that they account for half of all of the empirical use of NOS measures over the last six decades. This comparison reveals the dramatic differences in the type and scope of conceptions targeted by these instruments. While TOUS focuses almost exclusively on *Characteristics of scientific conceptions* and *Non-NOS* items, VOSTS include substantially more emphasis on *Epistemic virtues and vices* and various social processes of science, and the VNOS includes a larger focus on *Scientific justification* and *Follow-up* items. In addition, none of these instruments includes any significant focus on scientific aims or sources.

Discussion

There are several ways of interpreting these findings. First, the historical and disciplinary context of the field of NOS inquiry is useful for interpreting the patterns observed. Second, specific instances of significant historical change are evaluable as more or less productive for understanding NOS-related cognition. Third, the analytic framework developed provides a snapshot of the kinds and scope of the conceptions targeted by specific NOS instruments, allowing for their critical comparison.

The historical progression in NOS conceptions targeted by the instruments surveyed reveals a complex array of change from the 1950s to the present. While the field has experienced dramatic changes in what researchers regard as crucial targets of inquiry, some issues have also endured over the history of the field. Within each of the categories of the framework, shifts in the focus of research accompany the development of a profusion of new types of constructs. This has occurred as the blossoming of the field of NOS research in the last two decades saw a myriad of new studies targeting increasingly richer arrays of science conceptions.

One important source of this change is in response to refinements in researchers' conception of the nature of NOS constructs. Abd-El-Khalik (2013) has described how early items instruments "…bundled the assessment of cognitive, affective, and attitudinal outcomes related to the nature of the scientific enterprise…" (p. 5). The current study excluded wholly affective and attitudinal instruments, and focused entirely on instruments targeting declarative beliefs. Affective or attitudinal items that did feature in the sampled surveys received the *Non-NOS* code. While the data does support Abd-El-Khalik's claim of an "unbundling" of affective, attitudinal and conceptual research targets by the 1970s, the very large-scale surveys of the 1980s onward reincorporated a significant amount of non-doxastic, attitudinal measures.

Figure 5 also supports Abd-El-Khalik's (2013) observation of the dramatic growth of open-ended and qualitative over forced-choice instruments from 1990 through 2012. Follow-up items dramatically increased over this period and were almost entirely absent from early work. While early research instruments were largely theoretically driven, with prior normative theories of science playing a central role in their formulation, later instruments were increasingly developed empirically (e.g. VOSTS) and featured a growing number of semi-structured interviews coupled with open-ended follow-up items. This shift is very likely to have been in part a response to the major criticisms of the field, specifically that they paid insufficient attention to the structure and subtlety of participants' cognitions (Aikenhead et al., 1989; Mackay & White, 1974).

The history and philosophy of science provides another possible source of change in NOS instrumentation. For example, Abd-El-Khalick and Lederman have argued that: "Changes in conceptions of NOS have mirrored major shifts in focus and emphasis in the fields of philosophy, sociology, and history of science" (2000, p. 666). Although they have not attempted to trace these relationships in detail, they have suggested that several important shifts are common to both traditions. One important transition identified is in response to the "externalist" turn associated with Kuhn (1977), from a prior focus on developing a normative account and justification of science (e.g. Popper, 1959), to a focus on the context of discovery involving the social, psychological or cultural features of science (e.g. Barnes, 1974; Longino, 1990). This might suggest that the NOS instrumentation of later decades devote a greater proportion of attention to constructs targeting these kinds of conceptions.

The findings of this study provide mixed support for Abd-El-Khalick and Lederman's claim. For example, each decade of surveyed instruments devoted around 6% of all items to the sub-category *Social processes of science*, yet in the 1950s and 1980s, this reached 18.92% and 13.86% respectively (Figure 5). This pattern does not show a straightforwardly increasing focus on social dimensions of NOS over the relevant period, i.e. the alleged Kuhnian transition. The *Co-influence of social, economic, cultural, religious, ethnic, political and national values* code underwent considerable increase over the decades (from near 0% to around 2%), yet the 1950s saw an outsize focus on this issue at 4.73% of items (Figure 13). Although only two instruments were included from the 1950s, they appeared to have devoted a larger proportion of attention to the social dimensions of NOS than the instruments of later decades. Similarly, Figure 11 shows that the proportional focus on the *Epistemic virtues and vices* sub-category, capturing the psychological dimensions of science, decreased from 23.65% in the 1950s to 5.12% in the 2000s (though the 1980s briefly but strongly reversed this trend at 12.64%). Complicating this is the shift involving a substantive reduction in the number of non-epistemic descriptors of scientists in the *Epistemic virtues and vices* category. However, when factoring out non-epistemic descriptors, most decades include 4% to 6% of items focused on epistemic virtues and vices, although again the 1950s and 1980s are both closer to 13%.

Supporting Abd-El-Khalick and Lederman's claim is the increasing proportion of focus on the *Scientific disagreement* sub-category, from 0% in the 1950s to 4.69% in the 2010s (Figure 5). Similarly, the *Scientific change* sub-category increased steadily from 3.38% in the 1950s to 10.83% in the 2010s, and the *Rival versus cohering scientific conceptions* code rose from 0% in the 1960s to 3.61% of all items by the 2010s. Arguably, these dimensions of science are less characteristic of pre-Kuhnian philosophy of science (e.g. Carnap, 1937; Hempel, 1965). The very radical growth in focus on the *Sources of science* sub-category provides considerable support for claim of a Kuhnian transition in NOS instrumentation (Figure 5). This is especially so given the increased focus on psychological elements of scientific inquiry (e.g. perception), as well as on contexts of scientific discovery over that of justification.

Abd-El-Khalick and Lederman have also claimed that from the 1980s, NOS constructs increasingly focused on the theory-laden nature of observation and the importance of human creativity in science. These claims are strongly supported by the data, especially the scientific sources captured by the code *Creativity, imagination and intuition* and the code identifying scientists as *creative, imaginative and intuitive* (Figure 11). These constructs do not feature much in early NOS instruments and show substantial growth over subsequent decades, although the 1980s actually represent a low decade for both of these constructs. The consistent growth in the *Theory-observation interaction* code, from an initial proportion below 1% by the 1970s to an average around 2% by the 2010s also supports their account (Figure 10).

Abd-El-Khalick and Lederman have also described the history of NOS instrumentation as increasingly focused on the empirical and tentative nature of science. Again, the evidence provides support for this view. In NOS instruments before the 1980s, 0.64% and 1.44% of all items involve invocation of the *Epistemic stances* subcategory, yet this proportion averages around 2% over the next four decades (Figure 5). Reference to epistemic stances constitutes an increasing proportion of the *Sources, justification and epistemic stances* category (which itself undergoes considerable growth), revealing an increasing focus on questions of scientific certainty, conjecture and tentativeness (Figure 5). Also supporting their account of an increasingly empirically focused account of NOS is the sustained increase in the *Evidence: empirical, experimental and observational* and *Testability, tests and testing* codes (Figure 8 and 10).

Abd-El-Khalick and Lederman have also suggested that from the 1990s, researchers increasingly represented science as universally comprehensible, though subject to important limitations. In contrast, the data shows that early instruments actually devoted a far larger overall proportion of attention to the *Extents and limitations of science* code (Figure 14). Similarly, the *Intelligibility, difficulty and understandability of science*" code decreased from an early high of over one percent, to considerably below one percent in later decades (Figure 14).

In sum, the findings of this study only partly support the relatively monolithic kinds of changes proposed by Abd-El-Khalick and Lederman. Instead, the data presents a more mixed and nuanced pattern, which incorporates some aspects of the hypothesized Kuhnian transition, yet does not demonstrate others. The data also reveals the under-representation of some important research targets, and an occasional dramatic over-sampling of others. In particular, the focus of the *Aims and values of science* category demonstrates clear changes, which involved a shift from non-epistemic to epistemic aims (Figure 5). Although this is an important and valuable refinement of the construct, epistemic aims of science comprise the smallest category of the NOS conceptions studied, and so could be subject to further investigation. In addition, the expanded framework defended in Chinn et al. (2011) includes a small but important focus on epistemic responsibilities, which involve the responsibilities that a person might bear in virtue of their beliefs, knowledge and expertise. For example, scientists might have specific responsibilities to hold true beliefs about important questions in their fields of study, and to avoid certain kinds of error. However, none of the assessments surveyed addressed issues of epistemic responsibility, suggesting that this might be a valuable area of future research. Similarly, while a great deal of research has investigated beliefs about the relative value of science for achieving epistemic over non-epistemic goals, few items have explicitly asked students whether scientific knowledge is valuable for its own sake, independent of the goods that it might produce.

The data reveals some productive changes in the focus of NOS instrumentation. For one, the dramatically declining attention to non-epistemic features of the *Scientific aims and values* category mirrors the *Epistemic virtues and vices* category. This change involves the increasing refinement and improvement of the relevant NOS constructs, and by the 2000s, non-epistemic references have virtually disappeared from these categories (Figure 5). The considerable growth of items focused on the epistemic status and demarcation of science represents a valuable area of growth, given the importance of these conceptions for the acceptance of scientific knowledge in learners and the public (Figure 7). Similarly, although the proportional focus on issues of scientific progress declined over the decades, the dramatic increase in focus on the reasons and causes of scientific change shows that these constructs have improved in their sophistication (Figure 9). Merely knowing that people believe that science changes is far less interesting than the more detailed beliefs they might have about how and why that change occurs. As discussed above, the steady increase in issues involving scientific disagreement (Figure 5), as well as those involving rival versus cohering scientific conceptions (Figure 10) represent valuable changes in NOS conceptualization. Similarly, the increased focus on the co-influence between social, political and economic forces and science counts as another valuable shift.

Conclusion

The coding scheme developed and applied in this study provides a comprehensive tool for assessing the scope of the conceptual coverage of NOS instrumentation, both of specific assessments and of the field as a whole. It allows for fine-grained determination of the kinds of conceptual coverage of different NOS instruments, and to determine whether assessments aimed at comprehensiveness (e.g. VOSTS) really are so.

In sum, the results reveal a complex set of relationships and changes among a variegated and fine-grained set of analytic categories, over a long and important period of research. Further analysis using the this dataset is likely to be fruitful, especially in light of historical changes in the conceptions of science held in psychology and philosophy, variation in what counts as pedagogically valuable target conceptions amongst science educators and educational researchers, as well as shifting social and political norms and values. These influences have the potential to explain much of the change that the field has undergone.

Finally, the study demonstrates the value of the Chinn et al. (2011) framework for exploring the epistemic conceptions implicit in NOS research. By considering the full array of instruments over the history of the field of research, and using a comprehensive framework for NOS conceptions, the study reveals the rich and complex history of the field and raises prospects for improving our understanding of the target of NOS inquiry.

Tables

Table 1. Nature of Science assessment instruments

Key: WR: Written; INT: Interview; MC: Multiple choice; OE: Open-ended; U: Unknown

ID	Year	Assessment name	Developer(s)	Survey type and response type	Item no.
ORS	1954	[Opinions Related to Science]	Wilson, L. L.	WR: 2 pt. Likert	26
ATSSC	1959	Attitudes Towards Science and Scientific Careers	Allen, H.	WR: 5 pt. Likert	95
TOUS	1961	Test on Understanding Science	Klopfer, L., & Cooley, W.	WR: MC	60
MATS	1963	[Measuring Attitudes Toward Science]	Dutton, W. H., & Stephens, L.	WR: 1 pt. checklist	5
STT	1965	[Science Teaching & Testing]	Nedelsky, L.	WR: OE	7
SPI	1966	Science Process Inventory	Welch W. W. & Pella M. O.	WR: 2 pt. Likert	135
WISP	1967	Wisconsin Inventory of Science Processes	Scientific Literacy Research Group	WR: 3 pt. Likert	93
NSS	1967	Nature of Science Scale	Kimball, M. E.	WR: 2 pt. Likert	29
SSS	1968	Science Support Scale	Schwirian, P. M.	WR: 5 pt. Likert	40
SAI	1970	Scientific Attitude Inventory	Moore, R.W., & Sutman F. X.	WR: 4 pt. Likert	60
NOST	1975	Nature of Science Test	Billeh, V. Y., & Hasan, O. E.	WR: MC	10
TOSA	1976	Test of Scientific Attitudes	Kozlow, M. J., & Nay, M. A.	WR: MC	40
NSKS	1977	Nature of Scientific Knowledge Scale	Rubba, P. A.	WR: 5 pt. Likert	48
TOSRA	1978	Test of Science Related Attitudes	Fraser, B. J.	WR: 5 pt. Likert	70
COST	1981	Conceptions of Scientific Theories Test	Cotham, J., & Smith, E.	WR: 4 pt. Likert	50
LOS	1982	Language of Science	Ogunniyi, M. B.	WR: 6 pt. Likert	47
VOSTS	1989	Views on Science, Technology and Society	Aikenhead, G., Ryan, A., & Desautels, J.	WR: MC	114
USK	1989	[Understanding Scientific Knowledge]	Carey, S., Evans, R., Honda, M., & Unger, C.	INT	20
VNOS-A	1990	Views on the Nature of Science: A	Lederman & O'Malley	WR, INT: OE	7
VVV	1992	Views About Science and Scientists	Rampal, A.	WR: OE, MC	8
TBAS	1993	[Teachers' Beliefs About Science]	Pomeroy, D.	WR: 4 pt. Likert	50
PSE	1994	[Physics Students Epistemologies]	Roth, W. M., & Roychoudhury, A.	WR: OE	18
PSS	1997	[Philosophy of Science Scale]	Alters, B. J.	WR: 4 pt. Likert, OE	20

MSAS	1997	Mathematics and Science Attitude Survey.	Paciorek, E.	WR: 4 pt. Likert	62
PUS	1996	[Pupils' Understanding of Science]	Solomon, J., Scott, L., & Duveen, J.	WR, INT: MC	6
ANOS	1997	[Aspects of the Nature of Science]	Leach, J., Driver, R., Millar, R., & Scott, P.	INT	10
PF	1997	Nature of Science Survey Statements	Palmquist, B.C., & Finley, F.N.	WR, INT: OE	81
VASS	1997	Views About Science Survey	Halloun, I.	WR: 4 pt. Likert	50
EBST	1997	Epistemological Beliefs of Students and Teachers	Blanco, R., & Niaz, M.	WR: OE	4
SAI (II)	1997	Science Attitude Inventory II	Moore, R. W., & Foy, R. L. H.	WR: 5 pt. Likert	52
SEB	1998	Scientific Epistemological Beliefs	Tsai, C. C.	WR: 5 pt. Likert	16
VNOS-B	1998	Views of Nature of Science (B)	Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G.	WR, INT: OE	7
USI	1999	[Undergraduate Science Images]	Ryder, J., Leach, J., Driver, R.	INT: OE	19
PFSL	1999	Personal Frameworks for Science Learning	Hogan, K.	INT: OE	45
VANOS	1999	[Views about the Nature of Science]	Haidar, A. H.	WR: 7 pt. Likert	44
STSC	1999	[Student Teachers' Science Concepts]	Murcia, K., & Schibeci, R.	WR: TF, OE	22
EBAPS	1999	Epistemological Beliefs Assessment for Physical Science	White, B., Elby, A., Frederiksen, J., & Schwarz, C.,	WR: 5 pt. Likert, MC	30
TADS	2000	Theory and Data Survey	Leach, J., Millar, R., Ryder, J., & Séré, M. G.	WR: 5 pt. Likert	7
NOSI	2000	Nature of Science Interview	Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G.	INT	23
PINOS	2000	[Pupils Ideas on the Nature of Science]	Irwin, A. R.	WR, INT: OE	7
TSSI	2000	The Thinking About Science Instrument	Cobern, W. W.	WR: 5 pt. Likert	60
PSKA	2000	[Public Science Knowledge and Attitudes]	Bauer, M. W., Petkova, K., & Boyadjieva, P.	WR: 5 pt. Likert	20
SCS	2001	[Student Conceptions of Science]	Moss, D.M.	INT: OE	23
VNOS-C	2001	Views of Nature of Science (C)	Lederman, N. G., Schwartz, R. S., Abd-El-Khalick, F., & Bell, R. L.	WR, INT: OE	10
USVS	2001	[University Scientists' Views of Science]	Bianchini, J. A., Whitney, D. J., Breton, T. D., & Hilton-Brown, B. A.	WR, INT: 5 pt. Likert, OE	29
NSTQ	2001	Nature of Science and Technology Questionnaire	Tairab, H. H.	WR: MC	8
VNOS-E	2002	Views of Nature of Science (E)	Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S.	WR, INT: OE	8
BNOS	2002	[Beliefs in the Nature of Science]	Zeidler, D.L., Walker, K.A., Ackett, W.A., & Simmons, M.L.	WR, INT: OE	13
TASI	2002	Thinking About Science Instrument	Cobern, W. W., & Loving, C. C.	WR: 5 pt. Likert	35

SUMS	2002	Students Understanding of Models in Science	Treagust, D. F., Chittleborough, G., & Mamiala, T. L.	WR: 5 pt. Likert	27
SUNOS	2002	[Student's Understanding of Science]	Lin, H. S., & Chen, C. C.	WR: MC	11
POS	2003	Philosophy of Science	Gwimbi, E.M., & Monk, M.	WR: 11 pt. Likert	24
CARS	2003	Changes in Attitude About the Relevance of Science	Siegel, M. A., & Ranney, M. A.	WR: 6 pt. Likert	59
MVNOSB	2003	Modified Views of Nature of Science -B	Bell, R. L., Blair, L. M., Crawford, B. A., & Lederman, N. G.	WR, INT: OE	8
SEM	2004	[Students' Epistemologies of Models]	Gobert, J. D. & Pallant, A.	WR: OE	6
SVOS	2004	[Student Views on Science]	Kang, S., Scharman, L.C. & Noh, T.	WR: 5 pt. Likert, OE	5
VOSI-Sci	2004	Views of Scientific Inquiry	Schwartz, R. S.	WR, INT: OE	20
VNOS-Sci	2004	Views of Scientific Inquiry	Schwartz, R. S.	WR, INT: OE	10
EBSS	2004	[Epistemic Beliefs of Science Students]	Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D.	WR: 5 pt. Likert	26
SEV	2005	Scientific Epistemological Views	Tsai, C.C., & Liu, S. Y.	WR, INT: 5 pt. Likert, OE	19
SUSI	2005	Student Understanding of Scientific Inquiry	Liang, L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J.	WR: 5 pt. Likert, OE	55
MATOSS	2005	Modified Attitude Towards Organized Science Scale	Brossard, D., Lewenstein, B., & Bonney, R.	WR: 5 pt. Likert, MC	9
SUSSI	2006	Student Understanding of Science and Scientific Inquiry	Liang, L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J.	WR: 5 pt. Likert, OE	30
SES	2006	[Students' Epistemologies of Science]	Smith, C. L., & Wenk, L.	INT: OE	23
VOSE	2006	Views on Science and Education	Chen, S.	WR: 5 pt. Likert	15
NOSQA	2006	Nature of Science Questionnaire A	Khishfe, R. & Lederman, N. G.	WR, INT: OE	5
SVOSTS	2006	[Selections from the VOSTS]	Trumbull, Scarano, & Bonney	WR: MC	5
CLASS	2006	The Colorado Learning Attitudes about Science Survey	Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E.	WR: 5 pt. Likert	42
VONOS	2007	[View of the Nature of Science]	Ibáñez-Orcajo, T., & Marínez-Aznar, M	WR, INT: MC, OE	12
NOSQB	2007	Nature of Science Questionnaire B	Liu, S., & Lederman, N.G.	WR, INT: OE	13
ATS	2007	Attitudes Towards Science	Kind, P., Jones, K., & Barmby, P.	WR: 4 pt. Likert	45
INPECIP	2007	Inventory of Teachers' Pedagogical & Scientific Beliefs	Da-Silva, C., Mellado, V., Ruiz, C., & Porlán, R.	WR: 2 pt. Likert	56
NOSS	2008	Nature of Science Survey	Khishfe, R.	WR, INT: OE	8

SSNOS	2008	Seven Scales on the Nature of Science	Urhahne, D., Kremer, K., & Mayer, J.	WR: 5 pt. Likert	44
NSKAS	2011	Nature of Scientific Knowledge and Attitudes Survey	Young, N. J.	WR: 5 pt. Likert	50
MOSQ	2011	Myths of Science Questionnaire	Buaraphan, K.	WR: 3 pt. Likert	14
SBANOS	2012	Students Beliefs and Attitudes About the Nature of Science and Doing Science	Spady, D.	WR, INT: 4 pt. Likert, OE	21
STC	2012	Science Teaching and Constructivism	Annafo, Y.	WR: 4 pt. Likert	30
US	2012	[Understanding Science]	Park, H.	WR: 5 pt. Likert, OE	33
TBNOS	2013	[Teacher Beliefs about Science]	Belo, N. A. H.	WR, INT: OE	21
SINOS	2013	Students' Ideas about Nature of Science	Chen, S., Chang, W. H., Lieu, S. C., Kao, H. L., Huang, M. T., & Lin, S. F.	WR: 5 pt. Likert	54

Notes:

- If the survey designers assigned no name to their survey, a name was included in square brackets.

Table 2. Overview of Nature of Science coding categories

Sub-categories	Codes
A. Scientific aims and values	
Epistemic aims of science	Generic aims; Truth & revealing reality; Discovery; Knowledge, Understanding, explanation & description; Laws & theories; Mathematics; Systematizing experience & observation; Simplicity, complexity, scope, precision & comprehensiveness; Aims of experiments; Other aims.
Non-epistemic aims	General non-epistemic aims; religious, moral & spiritual aims.
Non-epistemic values	Moral value & science
B. Nature and structure of scientific knowledge	
Defining and demarcating science and the sciences	Generic definitions; Demarcation; Science & technology; Epistemic status; Science & the supernatural; The extent & limitations of science; Relations among scientific disciplines; Science & art; Analogies & metaphors; Body of knowledge vs. practices & skills; Science vs. religion.
Kinds of scientific conceptions	Generic conceptions; Laws; Theories; Hypotheses; Models; Distinguishing laws, theories, hypotheses & models.
Characteristics of scientific conceptions	Relation of conceptions to world; Discovery vs. invention; Simplicity vs. complexity; Testability, tests & testing; Consistency, coherence, comprehensive & paradigms; Dangers & benefits of science; Intelligibility, difficulty & understandability of science; Explanation & understanding; Relevance, importance, significance & usefulness of science; Objectivity, subjectivity, neutrality, bias & distortion; Inaccurate conceptions: role, utility & value; Other characteristics.
Scientific change	Generic change in science; Nature & kinds of change; Reasons & causes of change; Scientific progress.
C. Sources, justification and epistemic stances	
Sources of scientific knowledge	Generic sources; Perception, observation & sensation; Creativity, imagination, intuition & curiosity; Inference, reasoning & logic; Evidence: empirical, experimental & observational; Prior knowledge; Expertise & authority; Diverse sources; Self vs. other.
Epistemic stances	Epistemic stances
Scientific Justification	Topic-Specific justifications; Prediction & retrodiction; Anomalies; Theory-Observation interaction; Generic justification; Scientific vignettes.

Scientific disagreement	Generic disagreement vs. agreement; Rival vs. cohering conceptions & controversy; Reasons & sources of disagreement; Resolving disagreement & achieving consensus.
D. Epistemic virtues, vices and responsibilities of scientists	
Epistemic virtues and vices of scientists	Open-Minded, flexible & capable of belief revision; Logical, rational & rule-Following; Ethical, honest & intellectually honest; Curious; Creative, imaginative & intuitive; Emotions & emotional; Dedicated, hard-working, patient & determined; Biased vs. unbiased; Objective vs. subjective; Styles of cognition; Intelligent, knowledgeable, educated, skilled, trained & able; Believes the world is intelligible, lawful & ordered; Careful, cautious, precise, skeptical, systematic, & exact; Various other epistemic traits.
Other characteristics of scientists	Age, gender & race; Religious, non-Religious & anti-Religious; Philosophical orientation; Other descriptors of scientists.
E. Reliable processes for achieving epistemic aims	
Characterizing scientific processes	Generic processes & method; Unitary, multiplicit & method-less; Logical & orderly vs. illogical & disorderly; Necessity & benefits of method; Scientific inquiry; Scientific freedom.
Kinds of scientific processes	Confirmation vs. disconfirmation; Replication, verification & converging evidence; Questions, questioning & searching for answers; Criteria for the evaluation & selection of conceptions; Experimentation: definition, reasons & necessity; Experimentation: procedures & character; Discarding & replacing conceptions; Information: data, data collection & interpretation; Error, trial & error; Statistics & mathematics; Other processes or aspects of method.
Social and institutional processes of science	Scientific community: collaboration, competition, distribution, teamwork & loyalty; Publication: reputation, credibility, prestige, peer review, evaluation & citation; Communication: discussion, argumentation, dissemination & information-flow; Co-Influence of social, economic, cultural, religious, ethnic, Political & national values on science; Government, corporate & special-interest influence; Feminism & gender roles; Science funding; Public knowledge & support of science; Social power & decision-making; Other & generic "influences on science."
F. Non-NOS and Follow-Up items	
Non-NOS and Follow up items	Scientific learning & teaching; The role & usefulness of learning science; Affect & attitudes towards science; Validity of science learning; Science & me; Science content knowledge; Requests for reasons, justification, evidence, explanation, elaboration & examples.

Table 3. Scientific aims and values coding category

Key: A: Percentage of total items; B: Percentage within over-arching category; C: Percentage of surveys that contain code

Code	Description	Examples	A	B	C
Sub-category: Epistemic aims of science					
Generic aims	Generic aims or purposes of science.	"What do you think the goal of science is?" "…What do you think is the purpose of science generally?"	0.67	11.56	14.81
Truth and revealing reality	Aims of revealing truth, absolute truth, approx. truth, or reality.	"A scientist aims to discover the absolute truth" "The object of scientific activity is to reveal reality." Scientific findings always lead to final truths."	0.28	4.76	7.41
Discovery	Aims of discovering new phenomena.	"The goal of science is to discover new things in the world and the universe." "The chief reward in scientific work is the thrill of discovery."	0.20	3.40	6.17
Knowledge, understanding, explanation and description	Aims of the production of knowledge, understanding, explanations and descriptions.	"The goal of science is to build a better understanding of the world around us." "Scientific theories, principles and laws aim to correctly describe the world around us." "Science is an attempt to explain natural phenomena." "Scientists are always interested in better explanations of things." "The development of new ideas is a scientist's greatest source of satisfaction."	0.75	12.93	14.81
Laws and theories	Aims of capturing laws and/or theories.	"The value of science lies in its theoretical products" "Its value lies in its theoretical aspects."	0.20	3.40	6.17
Mathematics	Aims involving math or quantification.	"The ultimate goal of all science is to reduce observations and phenomena to a collection of mathematical relationships."	0.08	1.36	2.47
Systematize experience and observation	Aims of making sense of experience and observation.	"The goal of scientific theories is to classify a part of human experiences." "The purpose of science is to establish intellectual control over experience…"	0.16	2.72	4.94
Simplicity, complexity, scope, detail, precision, comprehensive	Aims of simplicity, complexity, scope, comprehensiveness, precision, accuracy, and detail of science.	"Science aims at ever-increasing comprehensiveness and simplifications using mathematics as a simple, precise method of stating relationships." "There is an effort in science to build as great a number of laws, theories and concepts as possible" "Science is constantly working toward more detailed and more complex knowledge."	0.24	4.08	6.17
Aims of experiments	Aims, goals and purposes of experimentation.	"Why do you think that scientists do experiments?" "Why carry out experiments?" "What is its goal? [of an experiment]"	0.59	10.20	16.05

Other aspects of aims	Other aims of science (e.g. one vs. many aims, problem solving, etc.).	“Do these goals differ in different disciplines?” “Do you think all scientific inquiry has the same purpose?” “Scientific theories and laws are primarily intended as tools for problem-solving.”	0.16	2.72	4.94
Sub-category: Non-epistemic aims					
General non-epistemic aims	Non-epistemic aims and practical applicability of science (e.g. helping people, technology, practical achievements).	“Science is a technology-developing activity. It is devoted to serving mankind. Its value lies in its practical uses.” “The one primary purpose of science in human society is to increase man’s control over nature and to increase his ability to use natural resources so as to make life more comfortable.” “Electronics are examples of the really valuable products of science.”	1.58	27.21	19.75
Religious, moral and spiritual aims	Aims related to religion and religious values, moral or spiritual ends.	“One important function of science is to demonstrate the wonder and orderliness of God’s universe.”	0.08	1.36	1.23
Sub-category: Non-epistemic values					
Moral value and science	The moral value of scientific research (e.g. good, bad, right, wrong, evil); ethical systems in science; the sanctity of nature, etc.	“It is incorrect to judge a piece of scientific knowledge as being good or bad” “The processes of science are divorced from moral and ethical considerations.” “Scientific ethics (i.e. system of morals) is concerned with, amongst other things, the possible harm that could result from scientific experiments”	0.83	14.29	11.11

Table 4. Nature and structure of scientific knowledge coding category

Key: A: Percentage of total items; B: Percentage within over-arching category; C: Percentage of surveys that contain code

Code	Description	Examples	A	B	C
Sub-category: Defining and demarcating science					
Generic definitions	General definitions of science.	"What is science?" "What is science all about?" "Science is an attitude towards life and the environment." "Do you consider this person's investigation to be scientific?" "What does it mean to study something scientifically?"	1.38	3.95	37.04
Demarcation	Demarcation of science from non-science, pseudoscience.	"How does [science] differ from non-science (e.g., religion and art)?" "Has (a) universal criteria(on) for demarcating science from non-science been found?" "Science and technology can NOT help people make legal decisions; for example, deciding if a person is guilty or not guilty in a court of law."	0.55	1.58	14.81
Science and technology	The relation between science and technology.	"Science and technology are closely related to each other…" "Science and technology impact each other." "Which one of the following best describes the relation between science and technology today? …"	0.40	1.13	7.41
Epistemic status	Scientific knowledge as better or worse than other knowledge (in epistemic terms).	"Scientific knowledge is different from other kinds of knowledge in that it has higher status." "Science is the ideal of knowledge in that it is a set of statements which are objective…" "Science is the best source of reliable knowledge" "Scientific knowledge is the truest form of knowledge."	0.83	2.37	14.81
Science and the supernatural	The supernatural, deities, miracles, intelligent design, etc.	"Science rests on an assumption that the natural world cannot be altered by a supernatural being (for example, a deity)." "Should supernatural causes be considered in science if empirical evidence points to such causes?"	0.28	0.79	7.41
The extent and limitations of science	What science can and cannot discover.	"Science deals with all problems and it can provide correct answers to all questions." "There are certain physical events in the universe which science can never explain." "Can scientific questions ever be answered fully and finally?"	1.42	4.07	24.69
Relations among scientific disciplines	The relations amongst the sub-disciplines science.	"Scientific fields such as chemistry and biology have fixed boundaries or borders" "Science disciplines differ from one another in what is studied, techniques used, and outcomes sought, but they share a common purpose and philosophy." "The various sciences contribute to a single body of knowledge"	0.99	2.82	14.81

Science and art	The relation between science, art and the arts.	“The actual work of scientists can be described as art.” “A scientific theory is similar to a work of art in that they both express creativity.” “There seems to be two kinds of people, those who understand the sciences and those who understand the arts (for example, literature, history, business, law)…”	0.47	1.36	12.35
Analogies and metaphors	Any analogy or metaphor involving science (not as art, technology).	“The course of scientific discovery resembles the process of reaching a difficult judicial decision.” “Scientists are essentially magicians, making two blades of grass appear where one grew before.” “Science is the shooting of a rocket to the moon.”	0.12	0.34	2.47
Body of knowledge, practices, skills	Science defined as knowledge, methods, skills or practices.	“The best definition of science would be an organized body of knowledge.” “Science is a body of knowledge.” “Science is an organization of our knowledge to help us learn about nature.”	0.20	0.56	4.94
Science vs. religion	Conflict between religion and science (not Aims or Virtues).	“Scientists have questioned many religious beliefs…” “Religion and science are almost always at odds with each other.” “We depend too much of science and not enough on faith.” “Scientific knowledge tends to erode spiritual values.”	0.63	1.81	8.64
Sub-category: Kinds of scientific conceptions					
Generic conceptions	Generic output of science.	“What kinds of ideas do scientists have?” “What are scientists' ideas about?”	0.32	0.90	8.64
Laws	Definitions, role and importance of laws.	“Scientific laws are only scientists’ best attempt to explain a part of nature.” “Scientists discover laws which tell us exactly what is going on in nature.”	1.03	2.94	12.35
Theories	Definitions, roles and importance of theories.	“What is a scientific theory?” “A theory is a hypothesis that has been proven to be correct.” “Theories are tools used to describe, explain, and predict scientific phenomena.” “[Science’s] Its value lies in its theoretical aspects.”	1.03	2.94	13.58
Hypotheses	Definitions, roles and importance of hypotheses.	“What is a hypothesis?” “Have you ever heard of the word hypothesis?” “A scientific hypothesis is the same things as a scientific fact.”	0.36	1.02	7.41
Models and simulations	Definitions, roles and importance of models.	“What are models?” “Models are explanatory tools.” “Which statement best describes scientific models…?” “A model can be a diagram or a picture, a map, graph or a photo.” “How would you describe what a model (in science) is to someone who didn't know this term? Give two examples of models.”	0.51	1.47	9.88
Distinguishing laws, theories, hypotheses and models	The interrelation between ideas, laws, theories and hypotheses, etc.).	“Is there a difference between theories and laws?” “In comparison to laws, theories have less evidence to support them.” “Laws are proven theories.” “Scientific ideas develop from hypotheses to theories, and finally, if they are good enough, to being scientific laws.”	1.70	4.86	25.93

Sub-category: Characteristics of scientific conceptions					
Relation of conceptions to world	The relation between conceptions and the world (e.g. truth, falsity, absolute truth, verisimilitude).	"There is no such thing as a true scientific theory." "Do you think science knows the right answers?" "How close do models have to be to what they model?" "It is likely that much of the scientific information we have today will be demonstrated to be inaccurate or inadequate in the future." "Scientific theories describe a real external world which is independent of human perception"	4.63	13.22	50.62
Discovery vs. invention	Conceptions as created/constructed vs. discovered; natural vs. human sources.	"Scientists discover theories, because the theories are there in nature and scientists just have to find them." "Scientists invent theories, because theory invention comes from the mind." "Were atoms discovered, or did scientists imagine them?" "Classification schemes are imposed upon nature by the scientists: they are not inherent in the materials classified."	1.82	5.20	25.93
Simplicity vs. complexity	Scientific conceptions as simple vs. complex (not as Aims).	"If a choice is to be made between two different scientific theories, both of which account for the observed facts, the more complex is chosen." "Good scientific theories explain observations well. But good theories are also simple rather than complex."	0.75	2.15	9.88
Testability, tests and testing	Science as subject to test (empirical, observational, experimental, etc.).	"Science is/is not testable." "Scientific knowledge need not be capable of experimental test." "A scientific theory is only true when it has been empirically tested and statistically significant proof has been provided." "Scientific laws, theories, and concepts are tested against reliable observations."	1.38	3.95	23.46
Consistency, coherence, comprehensive, paradigms	Coherence amongst conceptions, scope of science, or "paradigm(s)."	"Consistency among test results is not a requirement for the acceptance of scientific knowledge." "Theories are validated by their connection to other, generally accepted theories" "Theories fit within certain paradigms." "Scientific knowledge is specific as opposed to comprehensive."	0.51	1.47	7.41
Dangers and benefits of science	Harms and benefits of science (not as aims).	"Science is beneficial." "Science and technology solve many social problems, but science and technology also cause many of these problems." "Science and its inventions have caused more harm than good" "In the long run, man's lot will be improved by scientific knowledge."	3.13	8.93	23.46
Intelligibility, difficulty and understandability of science	The degree to which science can be understood or easily understood.	"Modern science is much too complicated for the average citizen to understand and appreciate." "Almost anyone can understand science if she/he studies it enough." "Most people are not able to understand science." "It is the nature of science to be intriguing and mysterious."	0.71	2.03	13.58

Explanation and understanding	Scientific conceptions as generating explanation and understanding.	"Science explains phenomena" "Science helps us to understand our environment." "A good theory may be accepted when it can be shown to explain things as well as another theory." "Scientific theories explain scientific laws." "Science is helpful in understanding today's world."	0.71	2.03	18.52
Relevance, importance, influence, significance and usefulness of science	The relevance, significance, importance or usefulness of science and scientific conceptions.	"Science is unrelated to life experience." "Science is useful for the problems of everyday life" "Science is very important in this scientific age in which we live." "To appreciate modern society fully, a person must understand the importance of science." "Science today receives too little serious attention in the mass media." "Understanding science is really important for people who design rockets, but not important for politicians."	1.19	3.39	14.81
Objectivity, subjectivity, neutrality, bias and distortion	The objectivity vs. subjectivity, neutrality and biases of science (not Virtue).	"Scientific evidence can be biased (i.e. distorted) in the way that data are interpreted, recorded, reported or selected." "Scientific neutrality has never really been achieved." "No form of knowledge – including science – can ever be completely objective." "Science is rational and objective."	0.55	1.58	9.88
Inaccurate conceptions: role, utility and value	The role of scientific superseded or inaccurate conceptions.	"Can wrong ideas in science ever be useful?" "Old theories are of no use to scientists." "A useful theory may not be correct, but it is the best idea scientists have been able to think up." "Does scientific knowledge become out of date?"	0.40	1.13	8.64
Other characteristics of science	Other characteristics of scientific conceptions (e.g. universally vs. locally valid; aesthetics).	"Newton's laws of motion … apply to physical objects that may be located: (a) anywhere in the universe. (b) in specific places of the universe." "In reaction to Einstein's equation … scientists said, "Such a beautifully elegant equation must be a true description of nature." This quotation shows that scientists assume their equations or ideas should match the elegance of nature."	0.32	0.90	6.17
Sub-category: Scientific Change					
Generic change in science	Scientific knowledge as changing vs. unchanging.	"Do scientists change their ideas?" "Even when scientific investigations are done correctly, the knowledge that scientists discover from those investigations may change in the future." "If science changes, why should we learn it?"	2.77	7.91	51.85
Nature and kinds of scientific change	The specific ways in which science changes (evolutionary, cumulative, etc.).	"Scientific knowledge is cumulative. It increases with increasing observation" "Scientific knowledge … also goes through jumps." "Once accepted, scientific knowledge may be slightly modified but not totally revised." "The development of scientific knowledge often involves the change of concepts."	0.87	2.49	19.75

Reasons and causes of scientific change	Accounts of why science or scientific knowledge changes.	“Old theories are replaced by new theories because…?” “Scientific knowledge changes … because the old knowledge is reinterpreted in light of new discoveries...” “Serendipity advances science” “Scientific theories have changed over time simply because experimental techniques have improved.”	2.25	6.44	38.27
Scientific progress	The improvement and advancement of scientific knowledge.	“[Scientific] knowledge is developmental” “Science is a self-correcting enterprise.” “All current knowledge may be superseded by future knowledge” “When scientists have a good explanation, they do not try to make it better.”	0.79	2.26	16.05

Table 5. Sources, justification and epistemic stances coding category

Key: A: Percentage of total items; B: Percentage within over-arching category; C: Percentage of surveys that contain code

Code	Description	Examples	A	B	C
Sub-category: Sources of scientific knowledge					
Generic sources	General, non-specific sources of science.	"Where do scientists get their ideas?" "Where do scientific ideas come from?" "Where do you go when you have questions about a scientific issue?" "Where does a scientist get a hypothesis?"	0.24	1.01	6.17
Perception, observation and sensation	Perception, sensation or observation as a source of science; direct observation vs. inference.	"Theories are based directly on observation, where observation is exactly what you see." "Observation of natural phenomena and experimentation is the basis of scientific explanation." "Scientific theorizing starts with observing the world around us in as thorough and open a way as possible." "Physicists say that electrons and protons exist in an atom because: (a) they have seen these particles in their actual form with some instruments. (b) they have made observations that may be attributed to such particles."	2.18	9.26	20.99
Creativity, imagination, intuition and curiosity	Creativity, curiosity, intuition, and imagination as scientific sources.	"Some accepted scientific knowledge comes from dreams and hunches" "Scientific theories are as much a result of imagination and intuition as inference from experimental results." "Science is/is not creative." "The fundamental driving force in science is curiosity concerning the physical universe."	1.07	4.55	19.75
Inference, reasoning and logic	Reasoning, logic, and inference as a source.	"Scientific knowledge relies heavily but not entirely, on observation, experimental evidence, rational arguments, and skepticism."	0.24	1.01	4.94
Evidence: empirical, experimental and observational	Empirical or experimental evidence as a source of science.	"A scientist evaluates scientific claims exclusively through empirical evidence." "Science knowledge is based on evidence..." "Science knowledge is based on evidence. Scientists should not have personal opinions." "Scientific knowledge comes from experiments only..."	0.75	3.20	14.81
Prior knowledge	Background knowledge, assumptions, literature or beliefs as a source.	"Scientists create theories based on prior knowledge, observation and logic." "When they investigate a particular event in the natural world, physicists decide what data they need to collect: (a) based on what they already know in physics. (b) after observing the event in all possible details." "The first inclination of a scientist is to try and integrate new knowledge into old knowledge."	0.67	2.86	8.64

Expertise and authority	The expertise, trustworthiness, “authority” or professionalism of scientists as a source.	“The basis of scientific explanation is in authority.” “Only scientists can think over scientific research questions” “A bit of information is considered scientific from physicists’ perspective: (a) when it has well-established merits regarding the natural world. (b) when it is offered by a group of trustworthy physicists.” “Valid scientific knowledge requires the acknowledgement of scientists in relevant fields.” “Good ideas in science can come from anybody, not just from scientists.”	0.75	3.20	14.81
Diverse sources	Non-standard, non-traditional sources.	“There is a significant amount of scientific knowledge in folklore and myth.” “How a scientific theory is generated is irrelevant to its usefulness.”	0.32	1.35	4.94
Self vs. other	The self (vs. others) as a source.	“In science, new ideas can develop from one’s own questions and experiments.” “Our foremost scientists are primarily concerned with their own thoughts and ideas.”	0.16	0.67	4.94
Sub-category: Epistemic stances					
Epistemic Stances	The certainty vs. tentativeness of science or scientists.	“[Are] scientists are certain about … the atom?” “To what extent [are] theories …conjectural?” “Scientific knowledge while durable has a tentative character.” “Even when making predictions based on accurate knowledge, scientists and engineers can tell us only what probably might happen. They cannot tell what will happen for certain.”	2.37	10.10	43.21
Sub-category: S**cientific justification**					
Topic-specific justifications	How do scientists know p (i.e. their evidence or justification)?	“How do scientists know that dinosaurs really existed?” “How do scientists know what an atom looks like what you have described or drawn?” “Do you think this person is also justified in concluding that natural selection shapes the teeth of animals to fit specific food resources?”	1.19	5.05	25.93
Prediction and retro-diction	The role and importance of scientific prediction.	“Good theories must explain and predict new phenomena.” “Scientific theories should explain additional observations that were not used in developing the theories in the first place” “The essential test of a scientific theory is its ability to correctly predict future events.”	0.75	3.20	16.05
Anomalies and falsification	Findings that conflict with existing theory.	“If a scientist does an experiment and the results are not as he or she expected, would the scientist consider this a bad result?” “An entire theory is falsified if subject to a single contradictory fact.”	0.71	3.03	12.35

Theory-observation interaction	The co-influence of observation and theory; “pure” untainted observation.	“Observations are theory laden.” “Scientists research activities will be affected by their existing theories” “Observation is influenced by theories scientists hold, because experimental procedures differ according to theories scientists hold, hence observation differs.” “Scientists rigorously attempt to eliminate human perspective from our picture of the world”	2.06	8.75	30.86
Generic justification	Generic “warrant” or “justification.”	“What warrants (if any) are drawn upon by young people to justify their belief or disbelief in theory?”	0.08	0.34	2.47
Scientific vignettes	Complex descriptions or vignettes.	“Scientists disagree about the issue of global warming. Some scientists say that humans are warming the planet by the continuous burning of fossil fuels. Another group of scientists say that the influence of humans is insignificant compared with the natural, which have determined the weather for so long”	4.31	18.35	34.57
Sub-category: Scientific disagreement					
Generic disagreement vs. agreement	Generic agreement, disagreement and discord in science.	“Do scientists ever disagree?” “Why do scientists disagree?” “Scientists share certain beliefs and attitudes about what they do and how they view their work.” “Scientists with similar background knowledge are trained to make similar observations of the same events.”	1.50	6.40	22.22
Rival vs. cohering conceptions and controversy	Alternative and conflicting conceptions, explanations or findings in science.	“One data set justifies only one scientific conclusion versus many conclusions” “How can we explain the different conclusions drawn from the same data?” “Can two different scientific explanations of the same phenomenon be good?” “If two scientists do the same research, can their results differ?” “Scientific knowledge is unambiguous: only one theory can be true.”	3.09	13.13	43.21
Reasons and sources of disagreement	Reasons for why there is agreement or disagreement in science.	“When scientists disagree on an issue (for example, whether or not low-level radiation is harmful), they disagree mostly because they do not have all the facts. Such scientific opinion has NOTHING to do with moral values (right or wrong conduct) or with personal motives (personal recognition, pleasing employers, or pleasing funding agencies).”	0.59	2.53	12.35
Resolving disagreement and achieving consensus	How disagreement in science is (or can be) resolved and consensus achieved.	“Is it possible to determine which of two disagreeing experts are right? How?” “It is always possible/not always possible to determine which of two competing explanations is more powerful” “How are conflicts resolved?” “… Scientists make this decision by consensus; that is, proposers of the theory must convince a large majority of fellow scientists to believe the new theory.”	0.47	2.02	14.81

Table 6. Epistemic virtues, vices and responsibilities coding category

Key: A: Percentage of total items; B: Percentage within over-arching category; C: Percentage of surveys that contain code

Code	Description	Examples	A	B	C
Sub-category: Epistemic virtues and vices of scientists					
Open-minded, flexible and capable of belief revision	Scientists as open to new ideas and as changeable, vs. close-minded (also gullibility).	"Science has a unique attribute of openness, both of mind and openness of the realm of investigation." "Scientists are willing to change their ideas and beliefs when confronted by new evidence" "… scientists … can be fooled by what they see on TV or read in newspapers."	0.95	8.66	14.81
Logical, rational and rule-following	Scientists as following rules, logic or as rational (or as illogical, irrational).	"The best scientists are always very open-minded, logical, unbiased and objective in their work. These personal characteristics are needed for doing the best science." "One important function of science is to teach people to be critical thinkers, not believing everything they are told." "It is necessary for scientists to be keenly aware of the rules which they follow and the tools they use in their pursuit of knowledge."	0.63	5.78	11.11
Ethical, honest and intellectually honest	Scientists as ethical, moral or professional (or as unethical), as well as honest or as intellectually honest (or dishonest).	"… most scientists behave professionally and ethically (i.e. in a moral and honest way)" "Scientists compete for research funds and for who will be the first to make a discovery. Sometimes fierce competition causes scientists to act in secrecy, lift ideas from other scientists, and lobby for money. In other words, sometimes scientists ignore the ideals of science (ideals such as sharing results, honesty, independence, etc.)."	0.32	2.89	7.41
Curious	Scientists as curious or as wanting to find things out.	"A scientist is someone who is curious." "Ideas about science experiments come from being curious and thinking about how things work." "Curiosity motivates scientists to make their discoveries."	0.20	1.81	6.17
Creative, imaginative and intuitive	Scientists as creative, imaginative (or unimaginative) or intuitive.	"When scientists are conducting scientific research, will they use their imagination?" "Scientific knowledge is not a product of human imagination."	2.29	20.94	32.10
Emotions and emotional	Scientists and emotions.	"Human emotion plays no part in the creation of scientific knowledge." "The most ideal form of scientific discovery is that in which scientists divorce themselves from their own personal and emotional involvement with the inquiry." "Scientists are more emotional than other people"	0.12	1.08	3.70

Dedicated, hard-working, patient and determined	Scientists as dedicated, hard-working, lazy, patient and determined, etc.	"A scientific work requires a dedication that excludes many aspects of the lives of people in other fields of work." "Scientific work requires long years of labor and self-discipline." "Scientists have practically no family life or social life because they need to be so deeply involved in their work."	0.20	1.81	4.94
Biased vs. unbiased	Scientists as biased or unbiased.	"…because a theory's content may be influenced by what a scientist **wants** to believe. Bias has an influence." "Scientific research is not influenced by society and culture because scientists are trained to conduct "pure," unbiased studies."	0.51	4.69	11.11
Objective vs. subjective	Scientists as objective vs. subjective, impartial vs. partial, neutral vs. partisan; etc.	"Scientists can make totally objective observations, which are not influenced by other factors." "The exactness and impartiality of the scientist in performing and reporting laboratory experiments is probably due in large part to the knowledge that his work will be examined by other competent workers rather than to the fact that scientists are more impartial and objective than other men."	0.91	8.30	16.05
Styles of cognition	Scientists thinking styles and attitudes (e.g. sequential, unorthodox, nonconformist, etc.).	"The process of scientific discovery often involves an ability to look at things in ways which are not commonly accepted." "Non-sequential thinking, i.e. taking conceptual leaps, is characteristic of many scientists." "A scientist might aptly be described as a nonconformist."	0.71	6.50	12.35
Intelligent, knowledgeable, educated, skilled, trained and able	Scientists as intelligent, knowledgeable, educated, skilled, trained and able (as compared to natural talent).	"Scientists as a group are more intelligent than those in other lines of work; such as law, business and farming." "The best scientists also need other personal traits such as imagination, intelligence and honesty." "A good solid grounding in basic scientific facts and inherited scientific knowledge is essential before young scientists can go on to make discoveries of their own."	0.91	8.30	12.35
Believe the world is intelligible, lawful and ordered	Scientists as believing that the nature is ordered and understandable.	"Scientists operate on the belief that the basic rules of the universe can be discovered by careful, systematic study." "A basic characteristic of science is faith in the susceptibility of the physical universe to human ordering and understanding."	0.55	5.05	7.41
Careful, cautious, precise, skeptical, systematic and exact	Scientists as careful, cautious, precise, exact, meticulous, and skeptical.	"[Scientists]… are going to try their best to observe precisely." "Science is a systematic way of thinking" "Scientists go overboard on demanding evidence before drawing conclusions." "A scientist should be skeptical of anything but his own work."	0.24	2.17	6.17
Generic traits and other epistemic descriptors	Any other (probable) epistemic descriptors.	"What are some characteristics of a scientist?" "What do you think of when you think of a typical scientist?"	0.08	0.72	1.23

Sub-category: Other characteristics of scientists					
Age, gender and race	Scientists' gender, race or age.	"The scientific community is mostly dominated by white men and is often unfriendly to minority people."	0.12	1.08	2.47
Religious, non-religious and anti-religious	Scientists as religious or pious (or as non-religious, atheist).	"A scientist's religious views will NOT make a difference to the scientific discoveries he or she makes." "A person can be both religious and scientific." "Scientists are against formal religion."	1.42	13.00	9.88
Other descriptors of scientists	Any other descriptors.	"Scientists are an odd lot." "Scientists are eggheads." "Scientists are communistic." "Scientists are usually unsociable." "There is no place in science for sexual deviants such as homosexuals."	0.12	1.08	3.70
Philosophical orientation	Scientists' philosophical views, including their ontology or the metaphysics (incl. positivism, realism, etc.)	"An ontological perspective with logical positivism is naïve." "Indicate your strength of belief in each of the following four basic philosophies as they relate to the epistemology of theories of the structure of space (geometry). For example, a priorism: 20%; conventionalism: 35%; positivism: 35%; realism: 10% (to total 100%)….)" "Those people who carry on the practices of science assume that… matter is an idea, not reality."	0.67	6.14	4.94

Table 7. Reliable processes for achieving epistemic aims coding category

Key: A: Percentage of total items; B: Percentage within over-arching category; C: Percentage of surveys that contain code

Code	Description	Examples	A	B	C
Sub-category: Characteristics of scientific processes and methods					
Generic processes and methods	Generic methods of science, or recipe of methodological steps, phases, components (e.g. induction, deduction, abduction, etc.).	"How do scientists do their work?" "What method should scientists use?" "How do scientists achieve their goals and answer questions?" "What sorts of things do scientists do that help them reach those goals?" "Science is essentially characterized by the methods and processes it uses." "The acquisition of new scientific knowledge moves from observation to hypothesis to testing to generalizing to theory." "While biologists use the deductive method to a problem, physicists work inductively."	3.09	10.96	33.33
Unitary, multiplicit vs. no methods	None, one or more methods of science (includes universal vs. context-bound methods).	"Is there a single universal scientific method, or many methods?" "Scientists follow the same step-by-step scientific method." "Scientists use several methods according to circumstances. The scientific method is only one of those methods." "Scientific method is a myth which is usually read into the story after it has been completed."	1.98	7.02	32.10
Logical and orderly vs. illogical and disorderly	Method as orderly, logical, hierarchical, determinate.	"Scientific method consists of fixed set of steps." "Method requires advance planning and rigor." "Scientific investigations follow definite approved procedures." "Scientists can adjust their method of inquiry in the middle of an investigation and still obtain valid results."	0.91	3.23	17.28
Necessity and benefits of method	The necessity and benefits of scientific methods.	"The best scientists are those who follow the steps of the scientific method." "Scientists are not compelled to use the traditional scientific method."	0.79	2.81	16.05
Scientific inquiry	Mention of scientific "inquiry."	"Scientists are most likely to achieve discovery by focusing selectively on the topic of inquiry." "How would you define "scientific inquiry" as it is conducted in your field of research?"	0.20	0.70	2.47
Scientific freedom	Scientists as free vs. constrained in their methods.	"The scientist will make his maximum contribution to society when he has freedom to work on problems which interest him" "Scientists should be free to explore all phases of man's life and the universe about him."	0.28	0.98	7.41

Sub-category: Kinds of scientific processes and methods					
Confirmation vs. disconfirmation	Confirmation or disconfirmation.	"Say a scientist is going to do an experiment to test his or her idea. Would a scientist do an experiment that might prove this idea is wrong?" "A scientist should attempt to disprove his own hypotheses."	0.12	0.42	3.70
Replication, verification and converging evidence	Checking, replicating and verifying results; criteria of repeatability; multiple sources and converging evidence.	"Science checks on its results" "A piece of scientific knowledge will be accepted if the evidence can be obtained by other investigators working under similar conditions." "Do you think science and scientists strive more to produce new knowledge, to verify existing knowledge, or both?" "Experiments using the same materials and procedures will have exactly the same results." "Good theories are based on the results of many different experiments."	0.99	3.51	19.75
Questions, questioning and searching for answers	Science as the asking of, and searching for answers to, questions.	"The scientific method involves the generation of new questions." "How do scientists answer their questions?" "Science is a search for findings." "An essential characteristic of a scientist is the ability to ask the right questions."	0.40	1.40	7.41
Criteria for the evaluation and selection of conceptions	The qualitative evaluation of science; theory choice criteria.	"How can good scientific work be distinguished from bad scientific work?" "In practice, choices between competing theories are made purely on the basis of experimental results." "It is not always possible to tell which is the most powerful of two competing theories, no matter how many data are available."	2.06	7.30	25.93
Experimentation: definition, reasons and necessity	The definition and purpose of experiments.	"What is an experiment?" "Do scientists do experiments?" "A scientist can obtain a direct answer to any simple question concerning nature by means of a carefully designed experiment."	1.27	4.49	23.46
Experimentation: procedures and character	The methods and character of experiments.	"How do scientists carry out experiments?" "How does a scientist decide what experiment to do?" "Do you think that scientists know what is going to happen before to do an experiment?"	1.58	5.62	25.93
Discarding and replacing conceptions	The discarding of old theories for new ones.	"The process of scientific discovery often involves purposeful discard of accepted theory." "Many of the scientific theories of the past have been discarded or modified as they have been found inadequate."	0.75	2.67	19.75
Information: data, data collection and interpretation	Information, data and its collection, interpretation and use by scientists.	"What does the word "data" mean in science?" "Scientific data must not be interpreted by the scientist." "Scientists use their imagination and creativity when they analyze and interpret data." "Scientists should spend almost all their time gathering information."	1.27	4.49	14.81

Error, trial and error	Mistakes or errors made in science, including trial and error.	"Are scientists ever wrong?" "Do scientists make errors? Why?" "Can scientists make mistakes or be wrong? How?" "Scientists should NOT make errors in their work because these errors slow the advance of science." "We accept scientific knowledge even though it may contain error."	0.91	3.23	13.58
Statistics, probability and mathematics	Statistics, probability, math and quantification as methods.	"Science without mathematics is impossible." "A…theory is only true when…statistically significant proof has been provided." "In physics, mathematical formulas express meaningful relationships among measurable quantities."	0.63	2.25	8.64
Other processes or aspects of method	Other processes or methods (e.g. luck, chance, serendipity).	"Physicists' findings about the natural world are … accidental, depending on physicists' luck." "The greatest scientists of the past have often made use of lucky guesses or "hunches"."	0.44	1.54	7.41
Sub-category: Social and institutional processes of science					
Scientific community	The role of community in science (e.g., cooperation, loyalty).	"Do scientists work alone?" "What do you see as the purpose of collaboration in science research? How do you think these collaborations happen or get started? Do you think collaboration is important?" "Science is a competitive enterprise."	2.14	7.58	28.40
Publication: reputation, credibility, prestige, peer review, evaluation and citation	Peer and institutional evaluation and legitimation (e.g. reputation, prestige, citations.)	"Scientists publish their discoveries in scientific journals. They do this mainly to achieve credibility in the eyes of other scientists and funding agencies; thus, helping their own careers to advance." "The winning of the esteem of his associates is one of the main incentives for the scientist." "A scientist's reputation can be important in judging his findings as the techniques he uses in his research."	0.87	3.09	14.81
Communication: discussion, argumentation, dissemination and information-flow	Discussion, argument and information sharing; information flow; public access.	"By sharing their ideas publicly, scientists build upon each other's work. Without this open communication, science would come to a standstill." "The discussion, debates, and result sharing in science community is one major factor facilitating the growth of scientific knowledge."	0.83	2.95	16.05
Co-influence of social, economic, cultural, religious, ethnic, political and national values	The influence of society, culture, and politics on science; includes "multi-culturalism."	"Scientific investigations are influence by socio-cultural values (e.g., current trends, values)." "Scientific knowledge is the same in different cultures" "Even though science is an activity carried out by many different people, science hardly ever reflects values and viewpoints related to society (e.g. views on women, political beliefs)"	2.49	8.85	33.33

Government, corporate and special-interest influence	The influence of government, regulation, special-interest groups and corporations.	“Within Canada there are groups of people who feel strongly in favour of or strongly against some research field. Science and technology projects are influenced by these special interest groups (such as environmentalists, religious organizations, and animal rights people).” “There is little need for the legal regulation of scientific research.”	0.91	3.23	13.58
Feminism and gender	The relative role of the genders in science (not epistemic virtue).	“When doing science or technology, a good female scientist would carry out the job basically in the same way as a good male scientist.” “Girls have very little mechanical aptitude, and therefore should not consider scientific careers.” “Science should remain a predominantly male profession.”	1.03	3.65	9.88
Science funding	The role of finance and funding in science.	“In order to improve the quality of living in Canada, it would be better to spend money on technological research RATHER THAN scientific research.” “Money spent on science is well worth spending.”	0.67	2.39	8.64
Public knowledge, interest and support of science	Public interest, understanding, and support of science.	“Public understanding of science would contribute nothing to the advancement of science.”	0.71	2.53	8.64
Social power and decision-making	Scientists as having power or ability to set public policy.	“Scientists possess too much power in our society” “…Since scientists are specialists in this field we should accept their judgment in matters of public policy rather than attempt to educate the public to make decisions on scientific matters.”	0.44	1.54	4.94
Other and generic “influences on science”	Other influences on science.	“What guides scientific research?” “…How do scientists decide what and how to investigate? Describe all the factors you think influence the work of scientists...” “Scientists must avoid being influenced by anything outside of "pure" science.” “Has philosophy influenced the development of science?”	0.44	1.54	12.35

Table 8. Non-NOS and *Follow-up* items coding category

Key: A: Percentage of total items; B: Percentage within over-arching category; C: Percentage of surveys that contain code

Code	Description	Examples	A	B	C
Scientific learning and teaching	How science is, or should be, learnt or taught.	"Science education should be more about the learning of scientific processes than the learning of scientific facts." "The science course in high school should investigate the definitions of and the relationships between hypothesis, theory and law"	23.38	100	60.49
The role and usefulness of learning science	The use of science learning or teaching.	"Has science helped you?" "…Biology, chemistry and physics are not practical for me. They emphasize theoretical and technical details that have little to do with my day-to-day world."			
Affect and attitudes	Science as interesting, exciting, boring (etc.).	"I think being a scientist would be exciting." "Scientific work is boring." "Scientific work is monotonous." "Scientific work would be too hard for me." "Scientists have to study too much." "I find science difficult."			
Validity of science learning	The validity of science learning or teaching.	"Results that pupils get from their experiments are as valid as anybody else's."			
Science literacy and content	The content of scientific theories.	"Do ghosts haunt old houses at night?" "How was the earth made?" "Can any metal be made into a magnet?" "How would you describe an atom?"			
Reasons, justification, evidence, explanation, elaboration, examples	Requests for further support, explanation, justification, evidence, elaboration or reasons.	"Why do you believe that?" "Give reasons for your answer" "Explain your answer" "Please give examples to illustrate your answer"	6.57	100	33.33

References

Abd-El-Khalick, F. (2001). Embedding Nature of Science Instruction in Preservice Elementary Science Courses: Abandoning Scientism, But... *Journal of Science Teacher Education*, 12(3), 215-233.

Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education*, 82(4), 417-436.

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education,* 22, 665-701.

Abd-El-Khalick, F., & Akerson, V. L. (2004). Learning as conceptual change: Factors mediating the development of preservice elementary teachers' views of nature of science. *Science Education*, 88, 785-810.

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics-Physics Education Research.* 2(1), 010101.

Aikenhead, G., Ryan, A., & Desautels, J. (1989). Monitoring student views on science-technology-society issues: The development of multiple-choice items. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95(3), 518-542.

Allen, H. (1959). *Attitudes of certain high school seniors toward science and scientific careers*. Bureau of Publications, Teachers College, Columbia University.

Alters, B. J. (1997). Whose nature of science? *Journal of Research in Science Teaching*, 34*,* 39–55.

Annafo, Y. (2012). *Junior high school science teachers' views about nature of science and science teaching*. (doctoral dissertation).

Barnes, B. (1974). *Scientific Knowledge and Sociological Theory*. Routledge.

Barufaldi, J.P.*,* Bethel, L.J.*,* & Lamb, W.G. (1977)*.* The effect of a science methods course on the philosophical view of science among elementary education majors*. Journal of Research in Science Teaching*, 14*,* 289–294*.*

Bauer, M. W., Petkova, K., & Boyadjieva, P. (2000). Public knowledge of and attitudes to science: Alternative measures that may end the "science war." *Science, Technology & Human Values*, 25(1), 30-51.

Bartholomew, H., Osborne, J., & Ratcliffe, M. (2004). Teaching students "ideas-about-science": Five dimensions of effective practice. *Science Education*, 88(5), 655-682.

Bell, R. L., Blair, L. M., Crawford, B. A., & Lederman, N. G. (2003). Just do it? Impact of a science apprenticeship program on high school students' understandings of the nature of science and scientific inquiry. *Journal of Research in Science Teaching*. 40(5), 487-509.

Belo, N. A. H. (2013). *Engaging students in the study of physics: an investigation of physics teachers' belief systems about teaching and learning physics* (Doctoral dissertation, Leiden University Graduate School of Teaching (ICLON), Faculty of Science, Leiden University).

Bianchini, J. A., Whitney, D. J., Breton, T. D., & Hilton-Brown, B. A. (2001). Toward inclusive science education: University scientists' views of students, instructional practices, and the nature of science. *Science Education*, 86, 42-78.

Billeh, V. Y., & Hasan, O. E. (1975). Factors influencing teachers' gain in understanding the nature of science. *Journal of Research in Science Teaching*, 12(3), 209-219.

Blanco, R.*,* & Niaz, M. (1997)*.* Epistemological beliefs of students and teachers about the nature of science: From 'baconian inductive ascent' to the 'irrelevance' of scientific laws*. Instructional Science*, 25, 203–231*.*

Boyd, R., Gasper, P., & Trout, J. D. (Eds.). (1991). *The philosophy of science*. The MIT Press.

Brossard, D., Lewenstein, B., & Bonney, R. (2005). Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education*, 27(9), 1099-1121.

Buaraphan, K. (2011). Pre-service physics teachers' conceptions of nature of science. *US-China Education Review*, 8(2), 137-148.

Buckland, L.A., & Chinn, C.A. (2015) Epistemology. *Encyclopedia of Science Education*. Gunstone, R. (Ed.) Springer Press.

Carey, S., Evans, R., Honda, M., & Unger, C. (1989). "An experiment is when you try it and see if it works": A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, 11*,* 514–529.

Carnap, R. (1937). Testability and meaning--continued. *Philosophy of science*, *4*(1), 1-40.

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw a scientist test. *Science Education*, 67(2), 255-265.

Chen, S. (2006). Developing an instrument to assess views on nature of science and attitudes toward teaching science. *Science Education*, 31, 1-31.

Chen, S., Chang, W. H., Lieu, S. C., Kao, H. L., Huang, M. T., & Lin, S. F. (2013). Development of an empirically based questionnaire to investigate young students' Ideas about nature of science. *Journal of Research in Science Teaching*, 50(4). 408 - 430.

Chinn, C.A., Buckland L.A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: arguments from philosophy and psychology, *Educational Psychologist*, 46, 141–167*.*

Chinn, C. A., Rinehart, R. W. & Buckland, L. A. (2014) Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In Rapp, D. N. & Braasch, J. (Eds.) *Processing Inaccurate Information*.

Cobern, W.W. (2000). The nature of science and the role of knowledge and belief. *Science and Education*, 9, 219-246.

Cobern, W. W., & Loving, C. C. (2002). An investigation of preservice elementary teachers' thinking about science. *Journal of Research in Science Teaching*, 39, 1016-1031.

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 186-204.

Cotham, J.*,* & Smith, E. (1981)*.* Development and validation of the conceptions of scientific theories test*. Journal of Research in Science Teaching*, 18*,* 387–396*.*

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham: Open University Press.

Duschl, R. A. (1985). Science education and philosophy of science: Twenty-five years of mutually exclusive development. *School Science and Mathematics*, 85, 541-555.

Duschl, R. A. (1988). Abandoning the scientistic legacy of science education. *Science Education*, 72. 51–62.

Dutton, W. H., & Stephens, L. (1963). Measuring attitudes toward science. *School Science and Mathematics*, 63(1), 43-49.

Elby, A., & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education*, 85, 554–567.

Fraser, B. J. (1978). Development of a test of science-related attitudes. *Science Education*, 62(4), 509-515.

Gobert, J. D. & Pallant, A. (2004). Fostering students' epistemologies of models via authentic model-based tasks. *Journal of Science Education and Technology*, 13, 7-22.

Guerra-Ramos, M. T. (2012). Teachers' ideas about the nature of science: A critical analysis of research approaches and their contribution to pedagogical practice. *Science & Education*, 21(5), 631-655.

Gwimbi, E.M., & Monk, M. (2003). Study of classroom practice and classroom contexts amongst senior high school biology teachers in Harare, Zimbabwe. *Science Education*, 87, 207–223.

Haidar, A. H. (1999). Emirates pre-service and in-service teachers' views about the nature of science. *International Journal of Science Education,* 21, 807-822.

Halloun, I. (1997). Views about science and physics achievement: The VASS story. In *AIP Conference Proceedings*, 399, 605.

Hammer, D. (1994a). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12, 151-183.
Hempel, C. (1965). Aspects of scientific explanation and other essays in the philosophy of science.
Hodson, D. (1988). Toward a philosophically more valid science curriculum. *Science Education*, 72(1), 19-40.
Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
Hogan, K. (1999). Relating students' personal frameworks for science learning to their cognition in collaborative contexts. *Science Education*, 83, 1–32.
Ibáñez-Orcajo, T., & Marínez Aznar, M. (2007). Solving problems in Genetics, part III: Change in the view of the nature of science. *International Journal of Science Education*, 29, 747-769.
Irwin, A. R. (2000). Historical case studies: Teaching the nature of science in context. *Science Education,* 84, 5-26.
Kang, S., Scharman, L.C. & Noh, T. (2004). Examining students' views on the nature of science: Results from Korean 6th, 8th and 10th graders. *Science Education*, 89, 314-44.
Kelly, G. J., & Duschl, R. A. (2002). Toward a research agenda for epistemological studies in science education. Paper presented at the Annual Meeting of NARST 2002, New Orleans, LA.
Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching*, 39, 551–578.
Khishfe, R. & Lederman, N. G . (2006). Teaching nature of science within a controversial topic. *Journal of Research in Science Teaching*, 43, 377-394.
Khishfe, R. (2008). The development of seventh graders' views of nature of science. *Journal of Research in Science Teaching*, 45, 470-496.
Kimball, M. E. (1967). Understanding the nature of science: A comparison of scientists and science teachers. *Journal of Research in Science Teaching*, 5(2), 110-120.
Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education*, 29(7), 871-893.
Klopfer, L., & Cooley, W. (1961). Test on understanding science, Form W. *Princeton, NJ: Educational Testing Services*.
Kozlow, M. J., & Nay, M. A. (1976). An approach to measuring scientific attitudes. *Science Education*, 60(2), 147-172.
Leach, J., Millar, R., Ryder, J., & Séré, M. G. (2000). Epistemological understanding in science learning: The consistency of representations across contexts. *Learning and Instruction*, 10(6), 497-527.
Leach, J., Driver, R., Millar, R., & Scott, P. (1997). A study of progression in learning about 'the nature of science': Issues of conceptualisation and methodology. *International Journal of Science Education*, 19, 147 – 166.
Lederman, I. L. (1985). Science and technology in Europe: A survey. *Science and Public Policy*, 12, 131-143.
Lederman, N. G., & O'Malley, M. (1990). Students' perceptions of tentativeness in science: Development, use, and sources of change. *Science Education*, 74, 225-239.
Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29, 331-359.
Lederman, N. G., Wade, P.D., & Bell, R.L. (1998). Assessing understanding of the nature of science: A historical perspective. In McComas, W. (Ed.), *The nature of science in science education: Rationales and strategies*. 331–350. Dordrecht, The Netherlands: Kluwer Academic.
Lederman, N. G., Wade, P.D., & Bell, R.L. (1998). Assessing the nature of science: What is the nature of our assessments? Science and Education, 7, 595–615.
Lederman, N. G., Abd-El-Khalick, F.S., Bell, R.L., Schwartz, R., & Akerson, V.L. (2001). Assessing the ''un-assessable'': Views of nature of science questionnaire (VNOS). Paper

presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MO.

Lederman, N. G., Schwartz, R. S., Abd-El-Khalick, F., & Bell, R. L. (2001). Pre-service teachers' understanding and teaching of nature of science: An intervention study. *Canadian Journal of Math, Science & Technology Education*, 1(2), 135-160.

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, 497-521.

Lederman, N. G. (2006). Research on Nature of Science: Reflections on the Past, Anticipations of the Future. Paper presented at the *Asia-Pacific Forum on Science Learning and Teaching*, 7, taken from: http://www.ied.edu.hk/apfslt/v7_issue1/foreword/index.htm.

Lederman, N.G. (2007). Nature of science: past, present, and future. In: S.K. Abell & N.G. Lederman (Eds.) *Handbook of Research on Science Education*, Lawrence Erlbaum, Mahwah, NJ, 831-879.

Liang, L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J. (2005). Student Understanding of Scientific Inquiry (SUSI): development and validation of an assessment instrument. In *Eighth International History, Philosophy, Sociology & Science Teaching Conference. University of Leeds, UK*.

Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J. (2006). Student Understanding of Science and Scientific Inquiry (SUSSI): revision and further validation of an assessment instrument. In *Annual Conference of the National Association for Research in Science Teaching (NARST)*.

Lin, H. S., & Chen, C. C. (2002). Promoting preservice chemistry teachers' understanding about the nature of science through history. *Journal of Research in Science Teaching*, 39(9), 773-792.

Liu, S. Y., & Lederman, N. G. (2007). Exploring prospective teachers' worldviews and conceptions of nature of science. *International Journal of Science Education*, 29(10), 1281-1307.

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press.

Mackay, L. D., & White, R. T. (1974). Development of an alternative to Likert scaling: Tests of perceptions of scientists and self (TOPOSS). Research in Science Education, 4(1), 131-139.

McComas, W. F., Clough, M. P.,& Almazroa, H. (2000). The role and character of the nature of science in science education. *In* W. F. McComas *(Ed.), The Nature of Science in Science Education*: *Rationales and Strategies*. The Netherlands: Kluwer Academic. 3–39.

McComas, W. F., Clough, M. P., & Almazroa, H. (2002). *The role and character of the nature of science in science education* (pp. 3-39). Springer Netherlands.

McComas, W. F., & Olson, J. K. (1998). The nature of science in international science education standards documents. In W. F. McComas (Ed.), *The nature of science in science education: Rationales and strategies*. 41 – 52. Dordrecht: Kluwer.

Moore, R.W., & Sutman, F. X. (1970). The development, field test and validation of an inventory of scientific attitudes. *Journal of Research in Science Teaching,* 7, 85-94.

Moore, R. W., & Foy, R. L. H. (1997). The scientific attitude inventory: A revision (SAI II). *Journal of Research in Science Teaching*, 34(4), 327-336.

Moss, D.M., Abrams, E.D., & Robb, J. (2001). Examining student conceptions of the nature of science. *International Journal of Science Education*, 8, 771–790.

Murcia, K., & Schibeci, R. (1999). Primary student teachers' conceptions of the nature of science. *International Journal of Science Education*, 21, 1123 – 1140.

Nedelsky, L. (1965). *Science teaching and testing*. Harcourt, Brace & World.

Nott, M., & Wellington, J. (1998). Eliciting, interpreting and developing teachers' understandings of the nature of science. *Science and Education, 7*, 579 – 594.

Ogunniyi, M. B. (1982). An analysis of prospective science teachers' understanding of the nature of science. *Journal of Research in Science Teaching*, 19(1), 25-32.

Osborne, J. (2002), Learning and Teaching About the Nature of Science, in S. Amos & R. Boohan (Eds.) *Teaching Science in Secondary Schools: Perspectives on Practice*, Routledge: London, 227–237.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. A. (2003). What “ideas-about-science” should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40, 692–720.

Paciorek, E. (1997). Mathematics and science attitude survey. Retrieved June 18, 2013, from Rochester Institute of Technology: http://oerl.sri.com/instruments/up/studsurv/instr129.html

Palmquist, B.C., & Finley, F.N. (1997). Preservice teachers’ views of the nature of science during a post baccalaureate science teaching program. *Journal of Research in Science Teaching*, 34, 595–615.

Park, H. (2012). *Understanding of the Nature of Science: A Comparative Study of Canadian and Korean Students* (Doctoral dissertation, University of Toronto).

Pomeroy, D. (1993). Implications of teachers' beliefs about the nature of science: Comparison of the beliefs of scientists, secondary science teachers, and elementary teachers. *Science Education*, 77(3), 261-278.

Popper, K. R. (1959). The logic of scientific discovery. *London: Hutchinson*.

Rampal, A. (1992). Images of science a study of school and scientists: Teachers' views. 1. Characteristics of scientists. *Science Education*, 76(4), 415-436.

Roth, W. M., & Roychoudhury, A. (1994). Physics students' epistemologies and views about knowing and learning. *Journal of Research in Science Teaching*, 31(1), 5-30.

Rubba, P. A., & Andersen, H. O. (1978). Development of an instrument to assess secondary school students understanding of the nature of scientific knowledge. *Science Education*, 62(4), 449-458.

Ryder, J., Leach, J., & Driver, R. (1999). Undergraduate science students’ images of science. *Journal of Research in Science Teaching*, 36(2), 201-219.

Sadler, T.D., Chambers, F.W., & Zeidler, D.L. (2004). Student conceptualisations of the nature of science in response to a socioscientific issue. *International Journal of Science Education*, 26, 387–409.

Sandoval, W. A. (2005). Understanding students’ practical epistemologies and their influence on learning through inquiry. *Science Education*, 89, 634–656.

Schibeci, R. A., & Sorensen, I. (1983). Elementary school children's perceptions of scientists. *School Science and Mathematics*, 83(1), 14-20.

Schwirian, P.M., (1968). On measuring attitudes toward Science. *Science Education*, 172-179.

Scientific Literacy Research Group. (1967). *Wisconsin inventory of science processes*.

Siegel, M. A., & Ranney, M. A. (2003). Developing the changes in attitude about the relevance of science (CARS) questionnaire and assessing two high school science classes. *Journal of Research in Science Teaching*, 40(8), 757-775.

Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-grade students’ epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction*, 18, 349–422.

Smith, C. L., & Wenk, L. (2006). Relations among three first-year college students’ epistemologies of science. *Journal of Research in Science Teaching,*43, 747-785.

Solomon, J., Scott, L., & Duveen, J. (1996). Large-scale exploration of pupils’ understanding of the nature of science. *Science Education*, 80, 493–508.

Spady, D. (2012). *Effects of Storytelling on Students' Beliefs and Attitudes about the Nature of Science and Doing Science* (Doctoral dissertation, Montana State University--Bozeman).

Schwartz, R. S. (2004). Epistemological views in authentic science practice: A cross-discipline comparison of scientists’ views of nature of science and scientific inquiry. Unpublished doctoral dissertation, Department of Science and Mathematics Education, Oregon State University, Corvallis, OR.

Tairab, H. H. (2001). How do pre-service and in-service science teachers view the nature of science and technology? *Research in Science and Technological Education*, 19(2), 235-250.

Tao, P. K. (2003). Eliciting and developing junior secondary students’ understanding of the nature of science through peer collaboration instruction in science stories. *International Journal of Science Education*, 25, 147–171.

Tobin, K. & McRobbie, C. (1997). Beliefs about the nature of science and the enacted science curriculum. *Science and Education*, 6, 355-371.

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of scientific models in learning science. *International Journal of Science Education*, 24, 357–368.

Trumbull, D. J., Scarano, G., & Bonney, R. (2006). Relations among two teachers’ practices and beliefs, conceptualizations of the nature of science, and their implementation of student independent inquiry projects. *International Journal of science education*, 28(14), 1717-1750.

Tsai, C. C. (1998). An analysis of scientific epistemological beliefs and learning orientations of Taiwanese eighth graders, *Science Education*, 82(4), 473-489.

Tsai, C. C., & Liu, S. Y. (2005). Developing a multi-dimensional instrument for assessing students' epistemological views toward science. *International Journal of Science Education*. 27, 1621–1638.

Urhahne, D., Kremer, K., & Mayer, J. (2008). Welches Verständnis haben Jugendliche von der Natur der Naturwissenschaften? Entwicklung und erste Schritte zur Validierung eines Fragebogens [What is adolescents’ understanding of the nature of science? Development and first steps to validation of a questionnaire]. *Unterrichtswissenschaft*, 36, 72-94.

Waters-Adams, S. (2006). The relationship between understanding of the nature of science and practice: The influence of teachers’ beliefs about education, teaching and learning. *International Journal of Science Education*, 28, 919–944.

White, B., Elby, A., Frederiksen, J., & Schwarz, C. (1999). The epistemological beliefs assessment for physical science. Presented at the American Education Research Association, Montreal (unpublished).

Welch, W. W., & Pella M. O. (1967). The development of an instrument for inventorying knowledge of the processes of science. *Journal of Research in Science Teaching*, 5, 64–68.

Wilson, L. (1954). A study of opinions related to the nature of science and its purpose in society. Science Education, 38(2), 159–164.

Windschitl, M. (2004). Folk theories of “inquiry”: How pre-service teachers reproduce the discourse and practices of an atheoretical scientific method. *Journal of Research in Science Teaching*, 41(5), 481–512.

Young, N. J. (2011). The effect of explicit instruction on student views of the nature of science and science related attitudes. Department of Science Education, Unpublished master’s thesis, California State University, Long Beach, CA.

Zeidler, D.L., Walker, K.A., Ackett, W.A., & Simmons, M.L. (2002). Tangled up in views: Beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education*, 86, 343–367.

CHAPTER II: EPISTEMIC COGNITION AND RELIABLE PROCESSES OF KNOWLEDGE PRODUCTION

Abstract

Chinn, Buckland, and Samarapungavan (2011) have developed an expanded framework for conceptualizing epistemic cognition for purposes of psychological research. This study investigates a particularly under-researched component of the Chinn et al. framework of expanded epistemic cognition: beliefs about the reliable processes used to generate knowledge. The models and measures that currently dominate this field of research do not adequately trace these important features of epistemic cognition. To investigate epistemic beliefs of this kind, 19 participating undergraduate students provided interview and written data while engaged in spontaneous and guided reasoning about various knowledge-generating processes. The resulting data revealed important kinds of variation amongst participants in the range of knowledge-generating processes they considered relevant. It also revealed variation in the sophistication and depth of participants' understanding of the processes involved in different forms of knowledge production. This study therefore demonstrates the utility of the expanded set of constructs of the Chinn et al. framework for advancing the field of research into the nature and character of epistemic cognition. It also reveals the importance of studying beliefs about the processes used to generate knowledge, as well as the conditions of reliability on those processes.

Keywords: epistemic cognition, personal epistemology, reliabilism

Introduction

A flourishing sub-discipline within psychology investigates epistemological features of cognition, or epistemic cognition. This field targets the beliefs and practices of learner, novice, and expert participants in laboratory or naturalistic settings. It aims to trace people's beliefs about what knowledge is, how to construct knowledge, and how to obtain and use that knowledge (Buckland & Chinn, 2015). Researchers in the field have developed a rough consensus regarding the constructs used to trace an array of valuable epistemic concepts, primarily involving the sources, structure, justification, and certainty of knowledge. Some, like Schommer (1990) and Hofer and Pintrich (1997), have conceptualized epistemic cognition in terms of a multidimensional set of beliefs about knowledge. Others, like Kuhn and Weinstock (2002), have understood epistemic cognition in terms of the development of epistemic sophistication through stage-like maturation. Nonetheless, these researchers each seek to explain and predict variation in learning outcomes and practices that arise from various modes of epistemic cognition. Of particular interest in this research are the distinctive beliefs about knowledge that predict learning success, as well as the kinds of interventions that encourage useful and accurate epistemic beliefs.

In spite of the promise of the field, it has to date largely investigated beliefs about a critically limited array of constructs. In particular, it has conceptualized the nature of epistemic cognition in terms of a relatively small, simple, and general set of target beliefs. The focus on the distinctive *characteristics* of knowledge has also meant that there has been little investigation of beliefs about the concrete means by which to create and locate knowledge. This restricted conceptualization has meant that existing models have difficulty in explaining real, complex cases. In particular, the focus on a few general dimensions of belief fails to recognize the importance of a complex web of causal beliefs about the processes implicated in knowledge production. We argue that attention to this web is vital for understanding real epistemic cognition.

The web of beliefs involving processes is crucial to epistemic cognition because the production of knowledge depends critically on a wide array of causal, cognitive, social, and institutional processes. For example, the trustworthiness of eyewitness testimony depends crucially on facts about the perceptual and cognitive processes involved in vision and memory.

Thus, for the testimony of an eyewitness to count as creditable knowledge, it must meet a range of conditions of reliability. Observations must occur while sober, under the right lighting conditions and relatively recently. These factors count as conditions of reliability on the processes underlying the production of this kind of perceptual knowledge. In general, the legitimacy of a knowledge claim, and thus its evaluation, depends on the processes involved in its generation, and crucially on the conditions of reliability for those processes. Epistemic competence involves knowledge of a myriad of processes used to produce knowledge, and the ability to critically evaluate their conditions of reliability. Beliefs about these processes and their reliability conditions are thus vital for understanding epistemic cognition. These beliefs thus represent a valuable target for the field.

In spite of the value of a wider conceptualization of epistemic cognition, basic questions in this area are still unanswered. For example, it is not clear whether people are typically aware of the processes implicated in knowledge production, and whether they can distinguish between different kinds of processes. It is also unclear how rich and detailed is their understanding of the conditions of reliability upon which these processes depend. Researchers have not investigated the strengths and weaknesses of different conceptualizations of knowledge production, nor have they investigated inter- and intra-individual variability in these dimensions of epistemic belief.

This study therefore investigates a vital yet under-researched aspect of epistemic cognition. It contributes to the field by investigating beliefs and understanding of the processes of knowledge production, as well as of the reliability conditions of these processes. It targets participant beliefs about the range of processes involved in knowledge production, the richness of beliefs about their conditions of reliability, the extent of individual and topic-dependent differences, as well as variation in spontaneous versus guided episodes of reasoning. We argue that the epistemic beliefs literature is limited in its account of how people reason about knowledge production; in response, we develop and test new constructs.

Next, we present the critical challenge to the field of epistemic cognition research developed in Chinn, Buckland, and Samarapungavan (2011), and briefly overview the epistemological tradition of reliabilism that motivates this challenge. We then argue for the importance of beliefs about reliable processes, and conclude the Introduction with the research aims and guiding questions of the study.

Challenging Epistemic Cognition Research

In a challenge to the prevailing orthodoxy in the field, Chinn et al. (2011) have presented a range of criticisms of popular models and measures. They have also developed an expanded framework that explores constructs beyond the traditional four dimensions of epistemic belief. This framework consists of five inter-related components: (a) epistemic aims and epistemic value; (b) the structure of knowledge and other epistemic achievements; (c) the sources and justification of knowledge and epistemic stances; (d) epistemic virtues and vices; and (e) reliable processes for achieving knowledge (see also Chinn, Rinehart and Buckland, 2014).

Chinn et al. (2011) further recommended the development of a more fine-grained, situated, and interconnected set of research constructs for the field, tracing beliefs about a larger range of beliefs that might feature in the construction or evaluation of knowledge. Instead of the current, relatively simplistic, focus on participants' abstract characterizations of knowledge (e.g. knowledge as simple versus complex), they have argued that the field might develop a more complex and inter-linked set of constructs for tracing variation and development. Existing conceptualizations in the field also tend towards a very low level of granularity; for example, dimensions of authority versus of experience currently serve to conceptualize beliefs about the sources of knowledge. Most real knowledge, however, obtains its epistemic legitimation from multiple interacting sources. For example, a scientist depends on both their own experience and on the testimonial authority and expertise of others in drawing conclusions and developing theories. Chinn et al. have claimed that the low-granularity of the conceptualizations implicit in orthodox models of epistemic cognition is unlikely to capture this complexity.

This study therefore advances the program of research Chinn et al. (2011) have proposed, by investigating a particularly under-researched component of the expanded framework – beliefs about reliable processes. The study also targets more complex, situated, and authentic kinds of beliefs about knowledge than is traditionally the case. Instead of eliciting the relatively abstract descriptions of knowledge characteristic of existing studies (e.g. knowledge as certain versus uncertain), or the coarse set of candidate sources of knowledge (e.g. the source of authority versus experience), this study focuses on tracing the more detailed and inter-related sets of beliefs about the ways in which knowledge is produced. It also assesses these beliefs as situated in a diverse range of complex, authentic knowledge-producing contexts.

Despite the relative paucity of research into beliefs about processes of knowledge production in the field of epistemic cognition, some existing research has targeted beliefs about specific knowledge-producing processes. However, researchers have typically not regarded these beliefs as a part of *epistemic* cognition, and so they have not integrated these findings with the wider field of study. This research includes the study of beliefs about memory, attention, collaborative decision-making, and eyewitness testimony. For example, a large survey by Simons and Chabris (2011) found widespread misconceptions about key cognitive processes concerning the power, objectivity, permanence, and reliability of memory. Perhaps the most intensively investigated area of beliefs about processes involves beliefs about the reliability of eyewitness testimony. Although widely regarded as a powerful source of testimonial knowledge derived from reliable processes of perception, a considerable body of research has cast doubt on the trustworthiness of eyewitness testimony. Studies by Tversky and Marsh (2000) and Loftus (2005) demonstrated that participants readily integrate false information into perceptual memories, and biased retelling serves to alter these narratives even further. In spite of these serious challenges to the reliability of eyewitness testimony, misconceptions abound. Simons and Chabris (2011) found that nearly a quarter of respondents consider the testimony of a single eyewitness as sufficient to ensure criminal conviction of a defendant. These widespread misconceptions should be cause for concern amongst epistemic cognition researchers, given the centrality of processes of perception and memory in almost all knowledge production.

In the field of epistemic cognition itself, some early studies did target important processes of knowledge production. For example, Schommer (1990) as well as Jehng, Johnson, and Anderson (1993) studied beliefs about the speed, order, complexity, and innateness of the processes involved in learning. However, subsequent work in the field (e.g. Hofer & Pintrich,

1997) explicitly abandoned the focus on beliefs about processes of learning, arguing that these beliefs do not properly target the core epistemic concepts of knowledge and justification. One implication of this move is that much epistemic cognition research fails to engage with what is arguably the most important of the set of processes by which one can gain knowledge: the processes of learning. The current study thus aims to help reverse this trend, by using the fifth component of the Chinn et al. (2011) framework to integrate these disparate strands of research, including beliefs about learning, memory, attention, eyewitness testimony, and social processes of knowledge production under the umbrella of epistemic cognition.

Epistemological Reliabilism as Resource for Psychology

Those studying epistemic cognition have often explicitly drawn on the conceptual resources of philosophy, and there have been repeated calls for a closer association between the two fields — see, for example, Greene, Azevedo and Torney-Purta (2008) as well as Murphy, Alexander, Greene and Edwards (2007). Chinn et al. have maintained that in spite of these calls, a great deal of productive work in contemporary epistemology has not yet had an impact on the field. In particular, they argued that work in epistemic cognition would benefit from greater contact with work in the tradition of epistemological reliabilism.

Reliabilist epistemologists, e.g. Nozick (1981), Goldman (1986, 1999), Kitcher (1993) and Dretske (2000), conceptualize knowledge and justification in terms of the 'truth-conduciveness' of the processes by which beliefs are generated. That is, a true belief counts as knowledge for a reliabilist insofar as it is the product of a reliable process; reliability on this account is a measure of the tendency of a belief-forming process to produce true rather than false beliefs. For reliabilists, to have knowledge requires that one form beliefs in the right kinds of causal, cognitive, and social ways. What differentiates the "right" from "wrong" kinds of ways, in this view, is the reliability of these processes in generating true beliefs (or a high ratio of true to false beliefs). A schematic of the reliabilist account of knowledge is roughly as follows: a person S knows a proposition p if, and only if, S believes that p, p is true, and S used a reliable belief-forming process to form the belief that p (Becker, 2009).

Reliabilism has proved to be a fertile approach to addressing some of the classic problems of epistemology. For instance, the traditional challenge in epistemology has been to develop a conceptual analysis of knowledge, and thus to specify its necessary and sufficient conditions. For most epistemologists, beliefs might be true as a matter of luck alone (e.g. a fortuitously accurate guess about the outcome of a random throw of the dice), and simply being true is not sufficient for a belief to count as knowledge. Similarly, as argued by Gettier (1963) and others, even a *justified* true belief is not sufficient for knowledge, given that simple counter-examples to this analysis can be found. For reliabilists, a particular true belief counts as knowledge only if formed through the kinds of processes that reliably (rather than accidentally) deliver truth. The focus on reliability is what distinguishes cases of mere accidental true belief from genuine cases of (non-accidentally true) knowledge. For the reliabilist, knowledge is in part dependent on the reliability of the perceptual and cognitive processes used in belief-formation.

The reliabilist program is thus a distinctively “naturalist” program in epistemology, focusing on the specific causal, perceptual, cognitive, social, and institutional processes that underlie knowledge production. In part due to the growth of reliabilism, contemporary epistemology has increasingly dissected the natural and social processes implicated in the construction, transmission, and use of knowledge, as well as the conditions upon which the reliability of these processes depend.

One study of epistemic cognition that has explicitly incorporated a focus on epistemological reliabilism is Hennessey, Murphy, and Kulikowich (2013). These researchers investigated teachers’ beliefs about the utility of alternate epistemic practices, characterizing these beliefs in terms of a three-part framework of foundationalism, coherentism, and reliabilism. While this represents a valuable infusion of contemporary epistemology into the field (see also Murphy, Alexander, Greene and Hennessey, 2011), it nonetheless involves an abstract, high-level and monolithic description of participant beliefs as “reliabilist.” This research does not yet trace fine-grained, situated beliefs about the reliability of specific knowledge-generating processes, nor does it examine how these beliefs affect knowledge construction and evaluation.

The Importance of Beliefs about Reliable Processes

The Chinn et al. framework has thus brought the conceptual toolkit of reliabilist epistemology to the psychological investigation of epistemic cognition. Making a case for this approach requires showing that understanding a person's beliefs about knowledge-producing processes is vital for explaining and predicting their decision-making and behavior. We maintain that understanding beliefs about reliable processes is particularly important for making sense of the ways in which people assert claims to knowledge and ascribe knowledge to others.

For example, consider three candidate cases of knowledge: the perceptual claims of an eyewitness, the testimony of an expert, and the output of a group of inquirers. While eyewitness testimony serves as a powerful source of evidence for a claim, the reliability of the perceptual and cognitive processes involved in the seeing, recalling and reporting of an event is hostage to a variety of further conditions. One might ask, for example, what were the relevant viewing conditions during the eye witnessing of the events; or how long has it been since the observation of the events in question. Similar considerations arise concerning the inferential knowledge expressed in expert testimony. Assessing the trustworthiness of the expert's claims requires taking into account a number of factors, including their qualifications and motivations. These represent the reliability conditions that determine whether the testimony counts as knowledge. Similarly, the effectiveness of a group undertaking collaborative inquiry in acquiring knowledge is hostage to the ways in which the group seeks out, processes, and communicates information. These are the conditions of reliability on social processes of knowledge production.

To evaluate the knowledge claims involved in each of these three scenarios, one must recognize and understand the knowledge-producing processes they comprise. In particular, one must be aware of the conditions upon which the reliability of these processes depends. Without this understanding, one cannot properly evaluate the epistemic legitimacy of the claims expressed in the scenarios; and one is therefore unlikely to dependably distinguish truth from falsity (or warranted from unwarranted beliefs). This means that an important part of the development of epistemic sophistication involves gaining an understanding of the conditions of reliability of knowledge-generating processes. Making sense of why someone rejects a wellfounded claim can thus require finding out if there are relevant processes which that they fail to recognize or actively discount. Alternatively, understanding why someone accepts a tendentious and unsubstantiated claim can require exposing their misconceptions about the conditions on particular knowledge-generating processes. One source of error is ignorance or misunderstanding of the processes involved in a knowledge claim. Alternatively, even an expert might fail to recognize violations of important conditions of reliability in complex cases.

We argue that because of the low granularity of their constructs, typically targeting only a few general beliefs about knowledge, orthodox models of epistemic cognition are insensitive to epistemic beliefs involving processes. For example, on existing models, an individual might reject the testimony of the expert in the second scenario because they regard the expert as lacking adequate *authority*. Alternatively, they might attribute knowledge in the eyewitness case because they regard perceptual knowledge as deriving its legitimation from the source of *experience*. However, these analyses may well miss the more complex beliefs that could be in play. The individual might in fact regard the expert as a legitimate source, yet also attribute bias to their view on the issue in question. Alternatively, they might regard eyewitness testimony as a questionable form of evidence, yet be willing to ascribe knowledge in this case because they believe that the observations occurred under the best possible conditions. Orthodox models do not trace the details of beliefs about knowledge-generating processes that are likely to influence an individual's evaluation of these knowledge claims. Tracing beliefs about reliable processes is therefore likely to be critical if we are to improve our understanding of epistemic cognition.

Research Aims and Questions

This study aims to explore participants' cognitions involving a variety of vital knowledge-generating processes, across multiple domains of reasoning. To achieve this aim, the study investigates the kinds of processes participants identify as important for various knowledge claims, as well as their understanding of the conditions of reliability on those processes. It also investigates the impact of a short period of reflection and discussion about processes of knowledge production pertaining to participants' judgments about authentic, complex cases involving knowledge attribution.

The central research questions investigated in the study are:

1. What kinds of reliability conditions on knowledge-generating processes can participants spontaneously identify as important?
2. How do participants vary in the conditions on reliability that they consider important for knowledge claims?
3. How does participants' understanding about processes vary across different epistemic contexts (e.g. from individual testimony to collaborative inquiry)?
4. Does explicit and deliberate reflection on knowledge-producing processes affect participants' judgments about complex cases?

Method

Design

To answer these research questions, participants completed two tasks: a verbal interview task and a written essay task. In the interview task, discussion served to elicit participants' beliefs about processes of knowledge production and the conditions of reliability on these processes. This discussion involved guided reflection on knowledge-relevant processes and conditions, with prompts designed to elicit participants' understanding. In the essay task, participants expressed their judgments about processes by responding in several pages of writing to complex articles.

Each of these tasks served to reveal the kinds of processes and conditions on processes that participants were able to articulate as important for knowledge production. The detailed follow-up questions of the interview task also aimed to expose their understanding (or misunderstanding) of these processes. These elements of the study thus help to answer the first three research questions. To answer the fourth research question of the study, the order of the interview and written measures was counterbalanced, producing two experimental conditions to which participants were randomly assigned (with groups balanced by sex). In condition A, participants first completed the essay task and then the interview task; condition B reversed this order. This design provides a way of measuring whether the interview task affected the number of knowledge-relevant processes participants were able to identify in the essay task. That is, it assesses whether the explicit reflection on processes in the interview task made a difference to participants' recognition of the knowledge-relevant processes featuring in the essays.

The study therefore implements a 2X2 analysis of variance, with a between-subjects design. The dependent variable comprises the number of conditions on processes that participants were able to identify as relevant to knowledge production. The independent variable comprises the two conditions created by the counterbalanced order of the interview and essay tasks. This design thus serves to reveal the degree to which the more easily articulable and explicit epistemic cognitions revealed in the interviews relate to the implicit epistemic cognition that guides participants' judgments during the more authentic and complex essay task.

Participants

Nineteen undergraduates (ten male and nine female) attending Rutgers University participated in the study. A flyer posted on campus in served to recruit participants, to which they responded via email or telephone. Participants ranged from 19 to 25 years of age, with an average age of 21.7 years. Ethnicity was determined through self-identification, with 37% of participants self-identifying as White, Caucasian or of European extraction; 21% Indian or South Asian; 16% as Asian, and the remaining 26% as Hispanic, African, and Mixed or not identified. Participants ranged from two to five years of university experience, with an average of 3.15 years, and reported a diverse range of major subjects of study, including: English, Philosophy, French, Linguistics, American Studies, Statistics, Psychology, Biomedical Science, Engineering, Animal Science, Education, Ecology, Pharmacy, Communication, Exercise Science, Fine Arts and Computer Science. Self-reported GPA scores ranged from 2.5 to 3.91, with an average of 3.2.

Procedure

Data collection occurred from 22 May until 6 June 2012. Participants read, discussed, and signed a consent form, and the interviewer briefly introduced the procedures involved in the study. Those in Condition A began by reading each of the two essay vignettes and then writing their short essay responses; they then read each interview vignette and responded to a set of verbal questions. Participants in Condition B completed the interview and essay tasks in the reverse order. All interview data was audio recorded, and both interview and written data transcribed.

Materials

The materials used to gather data consisted of a set of two interview vignettes and a set of two essays, each with an associated set of either verbal or written questions respectively. Each of the two interview vignettes (Appendices A and B) are between one to three paragraphs in length, consisting of a short descriptive narrative. The first, or "eyewitness vignette," involves the identification of an alleged criminal defendant; the second, the "commission vignette" involves the formation of an investigative commission of collaborating experts. The essays (Appendix C) were adapted from authentic news articles, and are much longer and more complex than the interview vignettes at around three pages. The medical essay involves a case of disagreement about the role of cholesterol in heart disease, and the judicial essay involves the conviction of a death-row inmate based on the testimony of a single eyewitness.

Crucially, each of the states of affairs described in the vignettes and essays involves inquiry into specific knowledge claims (e.g. the guilt of the death-row inmate, etc.). Each also incorporates different processes that reliably secure knowledge only when implemented under the right conditions. The relevant processes involve perception, cognition, memory, gathering physical and testimonial evidence, recording and testing information, and coordinating collaborative inquiry. The knowledge-generating circumstances described in all of the materials share the crucial characteristic of being highly sensitive to the conditions under which the various processes described were carried out. For example, knowing the particular circumstances under which an eyewitness viewed the alleged perpetrator, including observational conditions (lighting, distance, etc.), perceptual conditions (e.g. eyesight, etc.), as well as cognitive conditions (e.g. mental state of eyewitness, etc.) is vital for determining their testimonial trustworthiness. The same applies to features of the narratives that pertain to investigative or institutional processes. Sophisticated epistemic judgments about the knowledge claims expressed in the narratives will thus need to show sensitivity to facts about the conditions under which these processes occurred.

The interview questions (Appendices A and B) served to elicit participants' beliefs about the reliability conditions on various knowledge-generating processes. They comprise the "Spontaneous" responses and "Reflective" response phases. In the Spontaneous responses, participants first stated whether they felt they were able to make a judgment about each case, and explained their reasoning. They were then asked to make a judgment about the case (e.g. whether they could judge in favor of the defendant or the plaintiff in the eyewitness vignette), and express their level of confidence about that judgment. Participants then described as much of the additional information that they would need in order to make a considered and final judgment about the case described. All responses generated from the request for explanation, as well as the spontaneous explanations and elaborations that often accompanied the initial interview questions, were included as part of the Spontaneous responses.

The initial Spontaneous interview questions aimed to elicit the most explicit and articulable beliefs that participants held about the reliability conditions on the knowledge-generating processes that featured in the vignettes. They also aimed to reveal participants' beliefs that are subject to relatively effortless and spontaneous recall. In the subsequent Reflective response section, participants answered a more detailed set of questions about processes, and they explained the relevance of these processes to their overall judgments about the vignettes. While the Spontaneous responses allowed participants to identify those conditions on processes that were subject to immediate retrieval and articulation, the subsequent questions scaffolded participants' further thinking about processes, revealing those beliefs that are less easily accessed and expressed, and therefore more tacit.

The interviewer repeated the Spontaneous/Reflective question format for each of the knowledge-producing processes involved. Thus, for the eyewitness vignette, participants explained what they would need to know about the character of the plaintiff, the night of the crime, the events of the alleged eye witnessing, the plaintiff's general description of the culprit, the police lineup, and the character of the defendant. For the commission vignette, the interviewer asked participants what they would need to find out about the members, the formation of the commission, its operations, how it produces publications, and the sources it relies upon. Participants also stated what else they would need to find out about the situation (if anything) in order to make their final decision.

The essays (Appendix C) were adapted from contemporary news articles describing complex cases involving either expert disagreement about a scientific health issue or a criminal conviction based on eyewitness testimony. Researchers changed key identifying names (e.g. the names of the disagreeing research groups, the name, resident state, and description of the criminal convicted) in order to avoid eliciting participants' prior knowledge of the cases. The essay questions briefly asked participants to make a judgment about the essays, specifically what they could claim to know about the situations described. Although embedded in a broader and more complex social context, the essays explored the same kinds of knowledge-generating processes of the interview vignettes. Each provided participants with varied opportunities to reflect on the processes by which knowledge is generated (e.g. perception, judicial and scientific investigation, collaborative inquiry, etc.).

The combined Flesch-Kincaid reading grade level of the essays was 11.4, and they were therefore at an appropriate level of complexity for the university-level participants. The essays were selected because they explore real-life, thought-provoking and controversial topics, and so were considered likely to interest and engage participants. Real news articles served to minimize the artificiality of the intervention, and to elicit the kinds of responses that participants might form in non-laboratory contexts. Selection of the medical essay (Appendix C) depended on its presentation of authentic and relevant biological theory and evidence, as well as description of conflicting claims of differing groups of experts. Understanding why the groups have come to come to hold their divergent views requires thinking about the differing knowledge-generating processes upon which they each rely. The selection of the judicial essay (Appendix C) was due the provocative nature of the death penalty issue, which is likely to engage and motivate participants. This essay content also starkly reveals the potentially dramatic implications of claims to perceptual knowledge – in this case involving a death sentence. In making a judgment about the likely guilt or innocence of the relevant party, participants needed to consider the legitimacy of the eyewitnesses' knowledge claims. They therefore needed to consider the reliability-conditions on the perceptual and cognitive processes that underlie this kind of knowledge.

The research materials engaged participants in reasoning about disparate forms of knowledge-making, from judgments about criminal guilt involving testimonial and forensic sources to judgments about the social and institutional basis of collaborative inquiry. In response to more and less structured episodes of questioning, participants described what they would need to find out in order to make their own knowledge claims about the circumstances described in the vignettes. That is, they considered what conditions were required for the processes implicated in the vignettes to count as good, reliable, knowledge-generating processes, and which conditions would undermine their acceptance of the resulting knowledge claims.

Data Analysis

The data generated in the study was analyzed using a coding scheme developed in a bottom-up, inductive manner, with each code identified from participants' responses to questions about vignettes and essays, and each of the overarching code categories developed from considering how the individual codes form groups. This section describes the coding scheme by presenting the major categories of codes, and briefly explaining each of the constituent codes themselves. Table 9 provides an overview of the 35 codes. These codes occur within seven over-arching categories, six of which capture the central processes implicated in much knowledge production. Each of the codes identifies a condition of reliability on these processes. Table 10 provides a brief description of each code as well as examples of participant discourse assigned that code. The code categories captured are: (1) Conditions on perceptual and observational processes; (2) Conditions on cognitive processes; (3) Conditions on processes of inquiry; (4) Conditions on social and testimonial processes; (5) Conditions on institutional processes; (6) Conditions on processes involving physical evidence, and (7) Other. The next section briefly explains each.

Conditions on perceptual and observational processes. This category identified discussion of perceptual and observational conditions of the knowledge-generating processes described in the vignettes. The first of the five individual codes, "Generic background conditions," includes any mention of generic, unspecified background conditions. The more specific "Spatial-temporal relations" code concerns the relations of space and time between the agents. The "Visual conditions" code captures any explicit reference to the visual conditions in play during observations. The "Visual acuity" code concerns the visual acuity and perceptual capacities of observers. The "Number of observers and observer accord" code captured reference to the number of distinct observers described, as well as their level of observational agreement.

Conditions on cognitive processes. This coding category targeted participants' discussion of the conditions on the cognitive processes of the agents described in the vignettes, particularly their beliefs, desires, and dispositions. The "Memory acuity and accuracy" code captures any mention of the possibility of problems with the agents' memory of events. The "Memory interval" code identifies mention of the amount of time that has passed since the events in question, with participants recognizing that "fading" of memory makes older memories less trustworthy. The "Confidence, commitment, and belief-stability" code captured any reference to levels of confidence, strength of belief and commitment, or the likelihood of changing one's mind. This code identified responses focused on how certain the agents were in their knowledge claims. The "Bias and corruption" code captured references to the possibility of cognitive bias relevant to evaluating the knowledge claims implicit in the vignettes. It identified any mention of prejudice that may have undermined the reliability of the agents' testimony. The "Cognitive diversity and multiple perspectives" code captured references to the role and value of diversity in the beliefs and perspectives of the collaborating inquirers. This code identified responses oriented on determining whether groups of inquirers displayed a suitably varied set of perspectives and experiences. The "Background knowledge" code captured reference to the prior knowledge of the inquirers, with high levels of knowledge (particularly scientific knowledge) in the relevant field counting as an important precondition for accepting the judgment of an inquirer on a particular topic. The “Mental State” code captured references to the cognitive clarity of the agents in terms of their levels of sobriety, stress, panic, tiredness, and mental health. Finally, the “Motive” code captured references to the motivational state of the agents, particularly with regard to the incentives that might exist for lying.

Conditions on processes of inquiry. This coding category captured participant responses oriented on the conditions underlying the methods and procedures followed in the vignettes. These codes share a common focus on how these conditions impact on the relevant knowledge attributions that result from this inquiry. The “Generic information sources” code was assigned to any general identification of the sources of information relied upon by inquirers. The “Scientific methods and sources” code captured reference to scientific research, publications, citations, statistics, or sources drawn upon. The “Physical description” and “Identity parade” codes identified the relevant aspects of the investigative procedures undertaken. The first focused on reference to the role of the physical description given by the eyewitness, while the second focused on the procedures undertaken to ensure a fair police lineup. The “Other inquiry methods” code captured references to generic and further references to the procedures followed by the agents in generating and securing their knowledge claims.

Conditions on social and testimonial processes. This coding category captured participants’ reference to conditions on the social and testimonial processes in the vignettes. These responses focused on the conditions underlying the interpersonal processes implicated in knowledge production, including the role of corroborating versus conflicting testimony, the assessment of the reliability of testimony through facts about history, character, and interpersonal conflict. These five codes thus share a common focus on social features of knowledge production, particularly processes involved in testimony and securing testimonial trust. The “Identity: character and background of sources” code identified responses that focused on the history and character of the testifying agents described. The “Identity: character and background of inquirers" code captured reference to the background of the inquirers themselves (e.g. the investigating officers, the inquiring scientists). The "Sanction: expertise and accreditation" code identified any mention of the epistemic authority of any of the agents. This typically involved discussion of their education, expertise, experience, and accreditation. The "Personal relations" code captured any mention of the personal relations that exist between the agents in the vignettes. Finally, the "Alibi" code captured references to any further testimony that might serve to corroborate the claims expressed in the vignettes.

Conditions on institutional processes. This coding category identified participant responses that referred to the ways in which the structure, formation, origins, and activities of institutions serve to help or hinder in the construction of knowledge. It therefore identified those utterances that were oriented on the conditions of the institutional processes implicated in the vignettes. The "Number of inquirers" code focused on the size of the commission. The "Collaborative structure," "Collaborative origins" and "Collaborative procedures" codes focus on the structure, procedures, and dynamics of the commission respectively. The "Procedures of argumentation" code focuses on the discussion, deliberation, debate, and conflict-resolution of the commission. The "Information flow and access" code captured responses oriented on the commission's sharing of information. The "Publication and communications" code captured discourse oriented on the ways in which the commission communicated its findings. Finally, the "Funding" code captured any mention of how the commission funded its investigations.

Conditions on evidential processes. This coding category captured references to the conditions on processes by which evidence is generated and evaluated. The "Forensic evidence" code identified specific physical forms of evidence that were relevant to evaluating the knowledge claims in the vignettes. The "Generic evidence evaluation" code identified any generic or global reference to the quality and quantity of the evidence in the vignettes. For example, it captured reference to the importance of conducting further investigations and research. The "Other evidence" code captured references to other kinds of evidence that did not necessarily track the knowledge-generating processes described in the vignettes, or simply to unspecified evidence.

Other. This coding category captured any further kinds of utterances, including participant misunderstandings of the tasks, the expression of their own personal view and judgment, the moral status of the issues explored in the vignettes, or any uninterpretable expressions. The 35 codes of the interview analysis, grouped into the seven overarching categories, thus trace a wide range of the reliability conditions of some important and ubiquitous knowledge-generating processes. The scheme therefore provides a valuable tool for understanding how these processes feature in participants' epistemic cognition.

Inter-Rater Reliability

A first coder, who was blind to participant and condition, initially assigned codes. A second coder then assigned codes to a randomized subset of participant responses, similarly blind to participant and condition. After practicing on a small sample of responses, the second coder assigned codes to 25% of participant responses to the interviews, achieving agreement with the first coder on 86% of the codes assigned. After a similar period of practice, the second coder coded 36% the essay codes, achieving agreement on 81% of codes assigned. The coding scheme therefore demonstrates moderate levels of inter-rater reliability.

Results

Descriptive Statistics

Participants expressed 640 distinct responses to the interview questions subject to analysis, and wrote 38 short essays. These responses exhibited considerable variation in the number of conditions on processes invoked, the kinds of processes they featured, and the kinds of reliability conditions that participants emphasized and understood. Figure 16 shows the number of proportionate number of codes assigned across all interviews, including both vignettes, as well as Spontaneous and Reflective phases. The total number of codes assigned to all the interview responses was 1174, with 1134 of those being the six process codes that involve conditions on reliable processes of knowledge production (i.e. excluding the "Other" code). The average number of process codes assigned was 59.7, the mode was 64, and the median was 58. The combined essay and interview data ranged from a low of 44 to a high of 91 process codes assigned, and some participants thus mentioned over twice the number of conditions on processes as other participants. Coders assigned an average of 1.77 process codes to participants' interview responses. They also assigned an equivalent number of process codes to participant responses elicited by the two vignettes: 583 process codes for the eyewitness vignette and 543 process codes for the commission vignette. The 38 essays received 146 codes in total, with 105 being specifically process codes. The essays received an average of 2.8 process codes, with a median of 5.0, a mode of 4.0 and a range from 2 to 10.

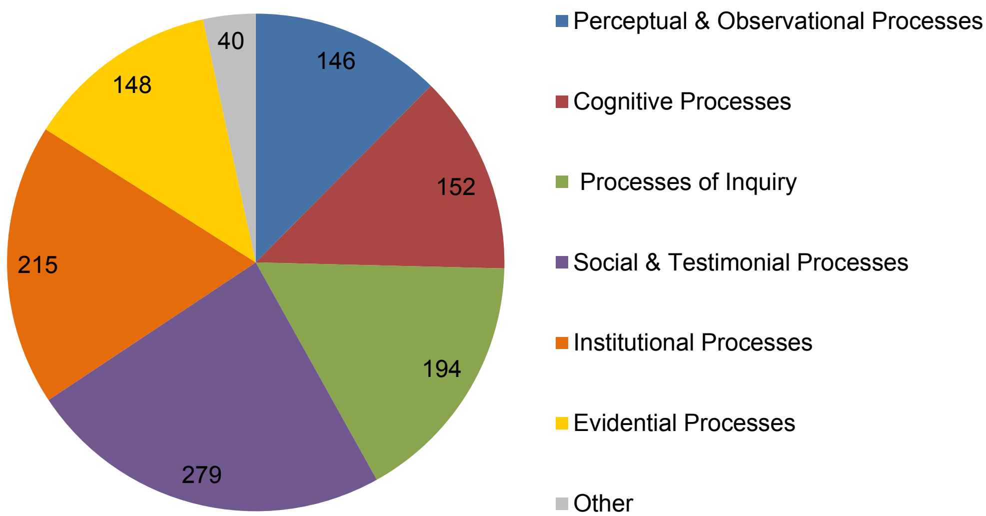


Figure 16. Number of interview codes: All responses. This figure shows the number of codes assigned to both interview vignettes and including both Spontaneous and Reflective responses.

The distribution of process codes also differed dramatically between the two interview vignettes. Table 9 shows that perceptual, inquiry, forensic and social processes dominated the responses to the eyewitness vignette, with a very small proportion focused on cognitive processes, and no attention at all directed at institutional processes. The eyewitness vignette thus predictably involved a focus on perceptual, inquiry and forensic processes, given the role of the eye witnessing of the crime and the subsequent police investigations. In contrast, the commission vignette responses featured no invocations of perceptual processes, and instead focused on the social and institutional features involved in the structure and formation of the inquiry commission.

Table 10 presents descriptive statistics showing the proportionate distribution of codes across sub-categories for the interviews. It shows that in the “Perception and observational processes” category, the “Spatial-temporal relations,” and “Visual conditions” codes dominate. A combination of the “Bias and corruption” code and the “Confidence, commitment, and belief stability” code dominates the “Cognitive processes” category, although the latter was more a feature of responses to the commission vignette rather than the eyewitness vignette, which did not feature much discussion of the level of confidence of the eyewitness. The table also shows that most of the “Processes of inquiry” category is devoted to the role of the “Physical description” and “Identity parade” codes. The “Generic information sources” code constitutes 20% of the category and the “Scientific methods and sources” code comprises another 15%. In the “Social and testimonial processes” category, the “Identity: character and background of sources” code constitutes 42% of the category, with the next largest being the “Sanction: expertise and accreditation” at 19%. The “Institutional processes” category, is dominated by the “Collaborative procedures” code, with the next largest being “Collaborative origins” at 21%, followed by “Publication and communications” at 21%.

Figures 17 and 18 show the distribution of codes assigned to the spontaneous interviews and essays of each participant, arranged from participants with the fewest to largest number of process codes assigned. In each figure, data bars distinguish the proportion of codes assigned to either of the two interview vignettes or the essays. Figures 17 and 18 shows the breakdown of process codes assigned to each participant, but provide proportional data showing the breakdown in the subcategories assigned. Figure 17 shows the proportional breakdown for the spontaneous interviews, whereas Figure 18 shows the proportional breakdown for the essays.

Inferential Statistics

We conducted a one-way analysis of variance on the number of process codes assigned to participants’ essays. This revealed that there was no statistically significant difference between the two experimental conditions of the study at the $p = 0.05$ level of significance ($F(1, 17) = 1.062$, $p = 0.317$). The initial engagement in inquiry about processes afforded by the interviews did not therefore measurably affect participants' subsequent performance on the essay task, at least in terms of the number of conditions on processes of knowledge-production mentioned

Analysis

Participants demonstrated considerable variation in their identification of processes in both the interview and essay responses. For example, 71% of the process codes mentioned by participant 12 were in response to the eyewitness vignette. In contrast, participant 11 showed a more even balance between processes mentioned in the interview vignettes, and mentioned significantly more processes than did other participants.

In addition to variation in the kinds of processes that participants identified, they also demonstrated considerable differences in their depth of understanding of these processes. One way in which participants demonstrated the sophistication of their conceptualization of processes of knowledge production was in the range of distinct conditions they spontaneously identified as relevant to their epistemic evaluation. Figure 17 clearly demonstrates this variation, in which participant 4 only considered perceptual and evidential processes, whereas participant 19 mentioned each kind of process. Similarly, Figure 18 shows participant 10, 4 and 2 as only mentioning two kinds of processes in the essays, whereas participant 19 again mentions all six.

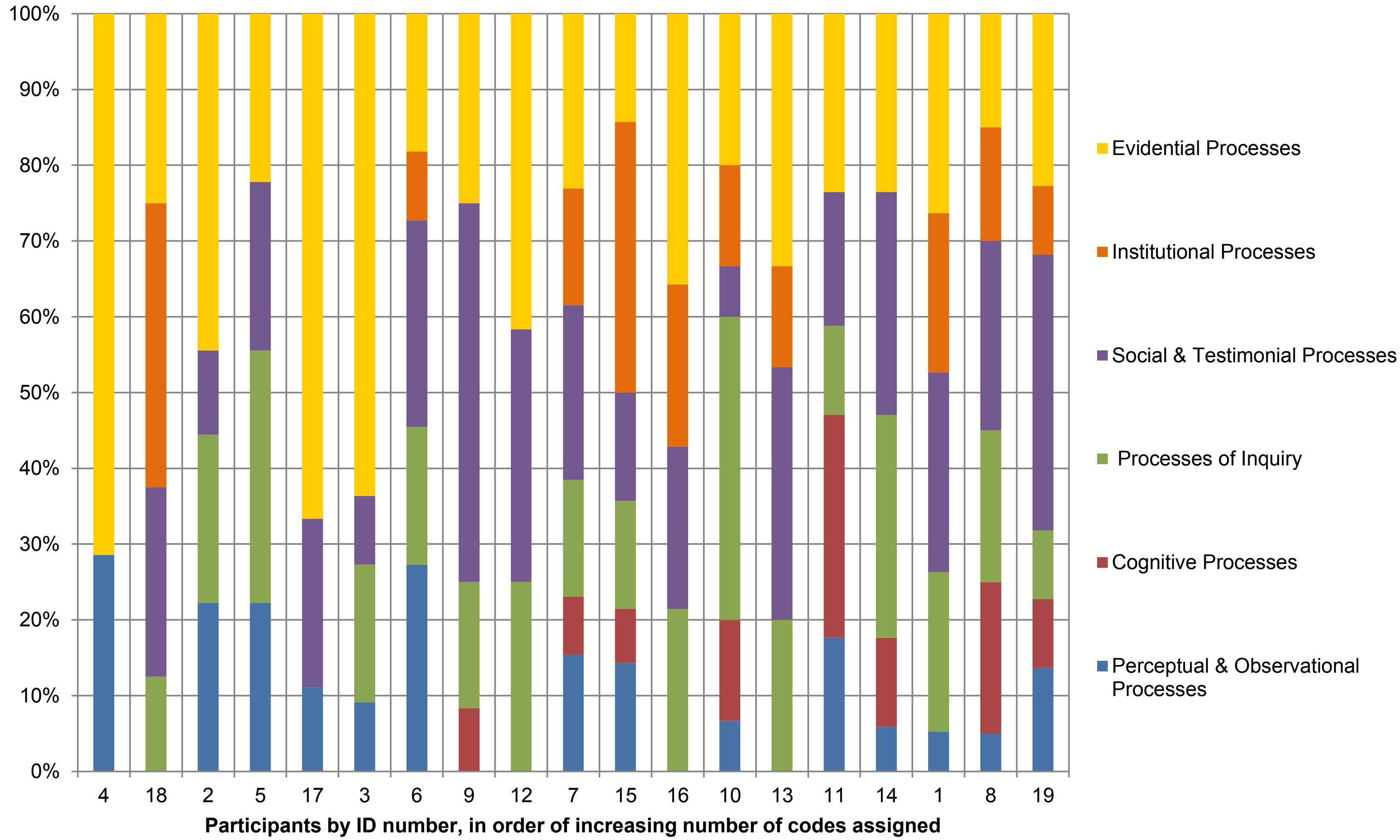


Figure 17. Proportion of process codes assigned by participant for spontaneous interviews. This figure shows the proportion of codes assigned to participants' interview responses, for both vignettes during spontaneous interview phases.

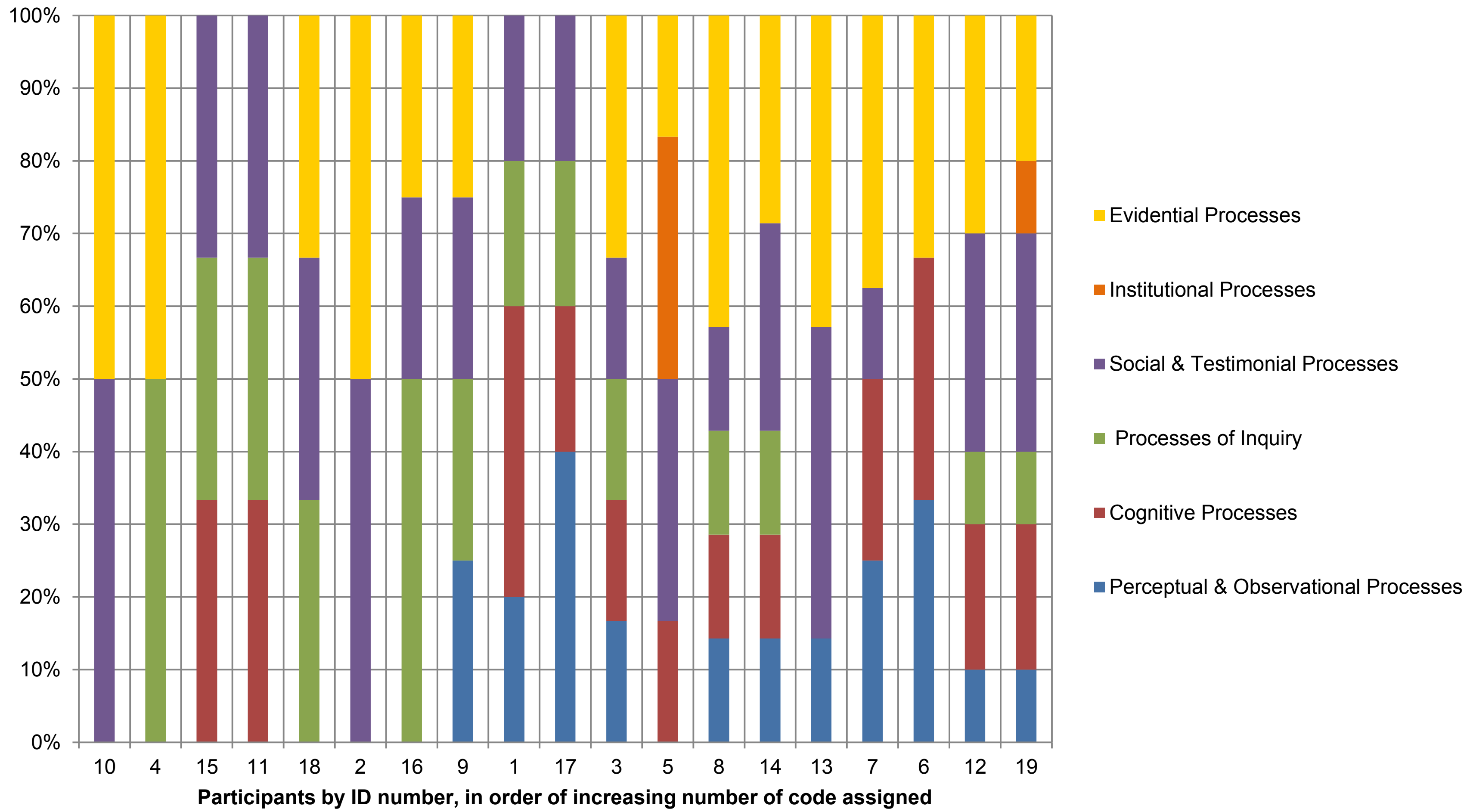


Figure 18. Proportion of process codes assigned by participant to essays. This figure shows the proportion of process codes assigned to participants' essay responses.

Figure 17 shows that all participants mentioned evidential processes of some kind during the spontaneous interview phase, with participants who mentioned the fewest distinct process kinds overall tending to focus on evidential processes more than did other participants. For example, over 70% of the spontaneous discussion of participant 4 focused on evidential processes, and participants 3 and 17 devoted over 60% to this kind of process. In contrast, four participants mentioned no evidential processes in the essays, with participants who mentioned the fewest number of conditions overall being more likely to include no reference to these processes. Figure 18 shows that only one participant mentioned each process kind and institutional processes were particularly under-represented, with only two participants considering them. The figures also shows that the seven participants at the lower end of the distribution mentioned no perceptual processes at all in their essays, whereas only one participant failed to mention social and testimonial processes across both interviews and essays.

Another way in which participants revealed variation in their epistemic sophistication was in their success in explaining the relevance of the conditions they identified to their overall evaluations. For example, in evaluating the perceptual knowledge described in the eyewitness vignette, participants generally identified multiple distinct conditions, including spatiotemporal relations, lighting conditions, and the number of observers. One participant stated that:

> “If he had a crystal clear picture of this guy then, he had a good chance that he is actually correct that the defendant stole money from him. If he barely saw him, he saw like a blur, it was really dark then basically it will narrow out the chance of the plaintiff being correct.”

In spite of this widespread competence in participants’ evaluations of perceptual knowledge claims, only one participant spontaneously identified the visual capabilities of the eyewitness as an important condition. This participant stated, “I want to know how the plaintiff’s eye sight is, if they need glasses, if they were wearing glasses at the time that they saw the person, if they have issues seeing at night.” Nonetheless, during the reflective phase of the interview, a further six participants mentioned visual competence, indicating that, at least with prompting, they did consider this to be an important condition for their evaluation of perceptual knowledge. Similarly, few participants spontaneously identified many conditions on cognitive processes as relevant to their evaluations of the eyewitness case, with only one participant focusing on issues involving the memory of the eyewitness. For the commission vignette, three participants spontaneously mentioned the confidence of the commission interviewees, and a further six considered issues of bias and corruption in evaluating their testimony. Only one participant spontaneously considered the question of cognitive diversity and multiple perspectives in their evaluation. In contrast, during the reflective phase, when specifically directed to focus on the cognitive processes involved in the vignettes, most participants mentioned multiple distinct conditions. In general, it was only during the reflective phase that participants focused on conditions on cognitive processes. For example, participant 12 exhibited a relatively sophisticated understanding of the importance of time for memory-dependent testimony, stating:

> "First of all, for how soon after the crime is because like everybody, over time the memory fades so it would be, if the plaintiff identified the defendant right after the crime, I feel that it would be more do to matching the actual description because it just happened and it's fresh in the plaintiff's mind."

Processes of inquiry and social and testimonial processes received the most attention from participants, both in the spontaneous and reflective phases. At least half of participants spontaneously mentioned the physical description of the suspect and the police lineup in the eyewitness vignette. Similarly, half of all participants spontaneously considered the sources of the information in the commission vignette, as well as the scientific character of their procedures. In the reflective phases, with explicit prompting, almost all participants considered these kinds of conditions. Some participants demonstrated a relatively sophisticated conception of the procedures of inquiry followed in the vignettes. For example, participant 18 clearly recognized the importance of the level of specificity in the description given to the police:

> "It would influence my decision because based on how specific or vague the description is, it could like just be pure chance that he guessed right or wrong cause there is not like a specific thing that he said that he mentioned beforehand about the burglar."

However, one area in which their understanding was sometimes limited was in the conditions of reliability on the police lineup described in the eyewitness vignette. While most participants tended to recognize that having wildly dissimilar individuals in the police lineup might be problematic, they generally battled to explain why this condition was important. For example, when pushed to explain why this aspect of the police lineup was important, participant 5 stated:

> "I don't know actually. I guess it would … influence it because I guess you know if whoever you know created the line-up maybe they pick criminals that best fit what the plaintiff is looking for or maybe they could pick people who are totally unrelated."

The social and testimonial processes codes also received a great deal of spontaneous attention from participants. The "Identity: character and background of sources" code was assigned more than any other code, and participants were generally very focused on determining the backgrounds of both the alleged eyewitness and the commission interviewees. For example, participant 5 was representative in stating that they needed to know:

> "The defendant's backgrounds, is he is a good person, do people say good things about him, do people say bad things about him, … would the defendant have a reason to take the money like would he even have some motive initially, like is he broke, you know, the defendant's whereabouts, where is he generally go at night, you know considering this crime took place at night, yeah I guess some of his habits and description of his personality and character."

Similarly, almost all participants spontaneously identified the issue of the sanction and accreditation of witnesses or those providing expert testimony. In particular, they focused on their qualifications in the relevant fields. Almost all participants raised the issue of prior personal relations between the eyewitness and the suspect, as well as between commission members and those commission interviewees. As one participant stated:

> "The first things I need to find out would be the relationship between the plaintiff and the defendant, is because if… the plaintiff and the defendant have an acquaintanceship, it could be anything from a misunderstanding to an actual burglary to the plaintiff allowing access into his home with somebody and that would still be a burglary but I'd think that, if you can prove that there's a relationship between the two for that point, then becomes a little bit more complicated, it becomes more he said, she said."

For this participant, the presence of prior personal relations between the vignette characters serves to introduce a further level of complexity to the situation, without automatically swaying them to one view over another. However, few participants showed a nuanced view of this kind. Many participants used the condition of prior personal relations between vignette characters as grounds for simply rejecting the knowledge claims made in the vignettes. For them, the presence of prior personal relations made it highly likely that the vignette characters were motivated by interpersonal animosity rather than by a concern for the truth. Complicating this picture was the fact that for many participants, prior personal relations between the eyewitness and the suspect in the eyewitness vignette served to improve the reliability of the eyewitness's knowledge claims. This is because participants judged that the process of identifying an individual already known to the eyewitness would be more reliable than their identification of a stranger.

Half of all participants spontaneously identified one or more institutional characteristics as important for their evaluation of the knowledge claims made in the commission vignette. In particular, their focus was largely on the collaborative and communicative procedures of the commission. No participants spontaneously considered the formation process of the commission, nor did they consider potential limitations on the information available to it. During the reflective phase, once the interviewer explicitly directed participants to consider institutional features, almost all participants demonstrated an awareness of the importance of these conditions. For example, participant 7 stated that would need to find out:

> "How the three commission leaders run the meetings, how long the meetings are, what they talk about, if they go out and do things that they talk about, like if they're just writing the report or if they're going out and investigating the genetically modified foods."

However, even during the reflective phase, only three participants considered issues involving the accessibility of information, with the majority of focus directed at the origins and procedures of the commission.

Discussion

Participants' thus considered a wide variety of processes involved in knowledge production in their deliberations, including perceptual, cognitive, social, testimonial, institutional, and evidential. Critical findings from the study comprise the considerable variation in participant performance in identifying processes, the distinctive conceptual limitations and affordances demonstrated by participants, the instructional implications and opportunities this reveals, as well as the ways in which the findings challenge the theoretical orthodoxy that currently dominates the field.

First, the study clearly revealed several different forms of variation in participants' epistemic cognition about processes of knowledge production. This variation was in terms of the number of distinct processes identified by participants, the number of times participants mentioned specific processes, the different array of processes identified across the various vignettes, as well as the difference in the pattern of participant responses across the spontaneous and reflective phases. Some participants mentioned conditions on processes at a significantly higher rate than did other participants, with the highest number being over twice the rate of the lowest. This implies that for some participants, a focus on processes and their associated conditions of reliability assumes a much greater role than for other participants. Figure 17 shows that the number of distinct *kinds* of reliability conditions spontaneously identified by participants (rather than the total number of tokens mentioned) also varied considerably. The highest scoring participant spontaneously mentioned four times as many distinct kinds of conditions on processes as did the lowest scoring participant. High-scoring participants thus exhibited a considerably richer conceptualization of knowledge production than did their peers. They therefore demonstrated a deeper understanding of the ways in which knowledge claims can achieve (or fail to achieve) their epistemic warrant, exhibiting a considerably greater set of resources for the evaluation of claims to knowledge. Existing models of epistemic cognition are unlikely to detect this considerable difference in epistemic resources. At best, they might distinguish the participants in terms of the sophistication of their general beliefs about the structure, sources, certainty, and justification of knowledge. Orthodox models would be insensitive to the more complex web of beliefs that guides peoples' reasoning about complex, authentic cases.

The data revealed that most participants tended to spontaneously identify a similar number of conditions on processes as compared across the two vignettes. However, those in the lower half of the distribution (and who thus mentioned fewer conditions overall), tended to identify more conditions in the eyewitness vignette. Participants 3 and 12, for example, mentioned around three times as many conditions on processes in the eyewitness vignette. This suggests that, at least for individuals who consider fewer of the relevant reliability-conditions, the perceptual and cognitive processes of the eyewitness vignette have greater spontaneous accessibility than the social, testimonial, and institutional processes of the commission vignette.

Both the interview and written data suggests that people demonstrate variation in the particular categories of reliability conditions they spontaneously consider when making knowledge judgments. Those who mentioned fewer conditions overall generally tended to mention fewer categories. For example, participant 4, who mentioned the lowest number of conditions overall, considered only perceptual and evidential processes. Five participants failed to mention perceptual processes at all, while nine overlooked all institutional processes. Only one participant made no mention of social and testimonial processes, and all participants devoted much of their spontaneous discussion phase to considering evidential processes. This supports the view that people vary considerably in their spontaneous identification both of conditions on processes, and on kinds of processes of knowledge production. It also suggests that people can exhibit very different patterns of processes that they regard as salient for knowledge evaluation.

In contrast to the spontaneous phase, guided discussion ensured that participants considered a wide range of knowledge-producing processes. In response to specific prompts, participants were generally able to recognize and evaluate these processes in discussion. While social and testimonial processes were a primary focus of participants, followed by institutional processes, considerable attention was devoted to all of the major kinds. This suggests that while ideas about processes might be difficult for many people to access spontaneously, directed reflection activates a complex web of beliefs about knowledge production.

The pattern of data observed also serves to reveal some of the distinctive misconceptions and weaknesses exhibited by participants in their conceptions knowledge production. Some of these may have implications for instruction and further research. The limited attention to the cognitive and institutional processes of the eyewitness vignette suggests that undergraduate conceptions of the epistemic basis of eyewitness testimony are limited in important ways. Similarly, while participants considering the commission vignette did spontaneously raise a number of important social and testimonial processes, they did not focus much of their attention on the way in which information flowed amongst collaborative inquirers, nor on the deliberative procedures governing their discussion and debate. In general, in the spontaneous phase, participants tended to overlook vital cognitive processes across all the presented narratives. Each of these patterns represents an important deficit in participants' epistemic conceptions, yet orthodox models of epistemic cognition are unlikely to be sensitive to them. Further research will help reveal the relationships between scores on orthodox measures and beliefs about processes. They will also help to show whether the distinctive pattern of weakness observed in participants generalizes to other populations.

The study identifies a range of weaknesses in epistemic cognition that process-oriented instruction might help alleviate. Instruction might also aim to inculcate learners with a greater sense of the importance of the full range of relevant processes for knowledge attributions. However, the inferential component of the analysis revealed that the brief period spent reflecting on processes of knowledge production did not detectably change participants' reference to processes in response to a complex reasoning task. Two important caveats apply. The first is that the limited number of participants in the study means that it is sensitive only to large differences between the groups. The second is that the low number of process codes elicited by the essays suggests that this kind of task is not particularly productive for eliciting cognitions about processes. Further research might explore more effective ways to elicit these cognitions, and to test these relationships with greater power. In spite of these caveats, the brief period of inquiry into reliable processes did not appear to generate the large impact on subsequent reasoning that the study would have been capable of detecting. Further research might reveal whether a longer, more structured instruction might more strongly affect subsequent process-oriented reasoning.

In sum, the study suggests that people do have a range of prior beliefs and assumptions about a variety of important knowledge-generating processes. It also shows that they differ in the range of processes they consider important, as well as the kinds of conditions they can call to mind when epistemically evaluating knowledge claims. The ability to recognize and evaluate processes of knowledge production represents an important determinant of epistemic expertise. It is likely to influence people's reliability in forming warranted and true beliefs in response to evidence. It is also likely that prior beliefs and intuitions about knowledge-producing processes is likely to influence their reasoning about the complex cases involving knowledge attribution encountered in daily life. Beliefs about processes therefore represent a valuable pedagogical as well as research target for the field. However, existing models of epistemic cognition simply not trace this important source of variation and sophistication in epistemic cognition; yet it is likely to be vital for explaining the strategies and outcomes demonstrated in real cognition.

Tables

Table 9. Overview of reliable processes of knowledge production codes

A: Interview on eyewitness vignette (Spontaneous): number of codes.
B: Interview on eyewitness vignette (Reflective): number of codes.
C: Interview on commission vignette (Spontaneous): number of codes.
D: Interview on commission vignette (Reflective): number of codes.
E: Written response to medical essay: number of codes.
F: Written response to judicial essay: number of codes.

Processes of Knowledge Production	Conditions on Process Reliability		A	B	C	D	E	F
Perceptual and observational	• Generic background • Spatial-temporal relations • Number of observers and observer accord	• Visual acuity • Visual conditions	25	121	0	0	15	0
Cognitive	• Memory acuity and accuracy • Memory interval • Confidence, commitment and belief-stability • Cognitive diversity and multiple perspectives	• Mental state • Motive • Background knowledge • Bias and corruption	5	53	13	79	14	4
Inquiry	• Physical description • Scientific methods and sources • Generic information sources	• Identity parade • Other inquiry methods	24	100	24	43	3	11
Social and testimonial	• Identity: character and background of sources • Identity: character and background of inquirers	• Alibi • Personal relations	26	119	35	98	21	5
Institutional	• Sanction: expertise and accreditation • Number of inquirers • Procedures of argumentation • Information flow and access • Publication and communication	• Funding • Collaborative structure • Collaborative origins • Collaborative procedures	0	0	27	186	0	3
Evidential	• Conduct further research or study • Generic evidence evaluation	• Forensic evidence • Other kinds of evidence	3	7	2	6	6	3
Other	• Other remarks		1	18	11	10	0	4

Table 10. Conditions on reliable processes of knowledge production codes

A: Interview process codes (percent within code category).

Conditions	Description	Examples	A
Conditions on perceptual and observational processes			
Generic background	Non-specific background conditions	"…what kind of background was at that time?"; "…What the plaintiff was doing, what the defendant was doing...."	8.2%
Spatial-temporal relations	Relative placement of the parties in space or time.	"…when the plaintiff saw the defendant climb over the wall did he get description of his face or was it from behind…"; "Where was the defendant on the night of the incident?"	36.3%
Visual conditions	Lighting or other visual conditions	"…or was there enough light that night because it happened at night so did he get a clear…"; "…but you don't actually get a good look at a person's face if there climbing over a wall."	28.1%
Visual acuity	Visual capacities of the observer	"That could help because if the plaintiff for example is partially blind, that could impair the vision of the plaintiff…"; "First I will see if his vision was like if he had a good vision…"	8.9%
Number of observers	Multiple and additional witnesses	"…or just like having one person see it makes lot less than having a bunch of people seeing it or like video evidence of the thinking documented…"; "…if there are any other witnesses saw him climb over the fence..."	18.5%
Conditions on cognitive processes			
Memory acuity & accuracy	Memory problems or ability; limitations or errors of memory	"I'd want to know the plaintiff's medical history if possible. Like it, how his, if he problems with like memory or retention or anything like that…"; "… see if his memory was like a decent memory, no faulty memory or really bad vision might have missed someone or sorts of details that weren't there…"	4.6%
Memory interval	Memory in relation to time	"Yes the longer, the more time that goes past, since the crime the more faulty the memory can be, or the more damaged it can be, or the more unreliable the memory…"	2.0%
Confidence, commitment & belief-stability	Level of confidence, strength of belief or changeability of mind	"…if the person change the story I would want to know why they change the story."; "Yes I think it would be better people with strong opinions because they keep some weight … "; "… if people are, serious about the topic…"	24.3%

Bias and corruption	Bias, prejudice, corruption, ulterior motives, agendas or undue influence involving the agents	"Whether he was black or white or any other racial to not have any kind of stereotype or bias against, uh, uh, the man…" ;"I'd want to know what the motives or any secondary motives that might exist here, like the commission leaders for example, I'd want to know any of their ulterior motives or their backgrounds or any relation they might have in general to genetically engineered food…"	28.3%
Cognitive diversity and multiple perspectives	Variety in cognitive states and dispositions (e.g. experiences, background)	"I'd also want to know the side of the groups and companies that support the genetically engineered food."; "Someone should have strong views. At least one or two people preferably two people with strong opposing views so that there isn't one person leading the group…"	9.2%
Background knowledge	Background knowledge as one of the conditions on the effectiveness of the committee	"If they're knowledgeable about uh these issues…"; "Because stupid people conducting a study would inevitably corrupt it to make it stupid. … Okay, I won't say they're stupid, I'm just saying that with no scientific experience the way they're talking about it wouldn't be valid because the information they know would not be able to contribute to it."	10.5%
Mental state	Cognitive clarity of the agents, (e.g. sobriety, stress, panic, tiredness, mental health)	"... if that, that person who witnessed was, uh, in his right mind, uh, he wasn't perhaps drunk or hallucinating..."; "I think when people are, when they are robbed they're in a state of panic, so you can easy see things that you don't think you actually see."	13.2%
Motive	Motivation of the agents	"… would the defendant have a reason to take the money like would he even have some motive initially, like is he broke, …"	7.9%
Conditions on processes of inquiry			
Generic information sources	Sources of information	"…where are they getting their information about genetically engineered foods from"; "…what sort of stuff they said…"; "What kind of data they are collecting?"	18.6%
Scientific methods and sources	Studies, research, publications, citations, statistics, scientists or scientific inquiry as a procedure or source	"is it based off whatever they've heard or have they been studying or doing research for long periods of time on cause and effect to make their decision?"; "…what the references they use are, if it's not just the interested members, if there's actual citations from papers."	16.0%
Physical description	Physical descriptions given by the agents	"First I will ask the plaintiff if he could describe his article of clothing and see if the defendant had matched up with those article of clothing is he still contain them…"; "How many people fit the description, mean are there any records in the police department already with people suiting that description?"	38.7%

Identity parade	Identity parades and criteria	"I guess who else was in the line-up like what their descriptions were, and whether they had limps or not…"; "I'd want to know in the police lineup how similar everyone looked to one another…"	22.7%
Other inquiry methods	Generic and other procedures and protocol of inquiring agents	"I would say because I'd want to know all the procedures they use and the protocols they use to acquire evidence and how long it take them to do everything...", "I mean, if the commission does the proper legwork then their results could be relied upon."	4.1%
Conditions on social and testimonial processes			
Identity: character & background of sources	Character, lifestyle, behavior of the agents (e.g. trustworthiness, criminal records, etc.)	"If the defendant had a pass history of burglaries of any type that of things he would steal and got caught with…"; "I would want to know a lot about the interested members of the public who were interviewed, such as the, like the backgrounds, like education, age, where they live, what they do for a living…"	39.4%
Identity: character & background of inquirers	Character and identity of the inquiring agents	"I'd want to know the backgrounds of every single individual involved in this commission, and the individuals that they interviewed."; "I would want to know the background. Age, education … history … I would want to know where they lived what their jobs were."	10.4%
Sanction: expertise & accreditation	Authority, accreditation, education or training of any of the inquiring agents	"I'd want to know how experienced, what their background on writing annual reports for publications are, that might be a hint as to their education level."; "The citizens, are they just regular people like or do they work like in a medical field or scientific field…"	19.4%
Personal relations	Prior history or relationship between the agents, or with other groups	"So why... is there a connection between them?"; "Did the plaintiff, does the plaintiff know the defendant, because it seems that he specifies this person specifically…"; "Basically who stands to gain and what alliances the members have and which groups and even if they didn't have any alliances with any group…"	16.5%
Alibi	Presence of an alibi	"Did he happen to have an alibi?"; "… if he was with anyone…."	14.3%
Conditions on institutional processes			
Number of inquirers	Relative size of the inquiring commission	"I don't remember which it is but if it's a small group it's easier to control information and its more prone to brainstorming phenomenon where basically the information is agreed upon by the head, the head authorities in the group would tend to spread around the group and reiterate it but if it's a large group it's hard to control a large fluctuating opinion of such many people."	4.2%

Collaborative structure	Structure of composition (e.g. “leaders”)	“Who are the leaders, what are those leaders’ backgrounds?”; “I would want to know if there’s like a head on the commission so if there’s someone who is in charge or if it’s just kind of a free for all for people to speak whenever they want to or if they’re told specifically when to speak.”	11.6%
Collaborative origins	Methods by which the inquiring group was formed	“I want to know how old the actual commission that was formed is like how many years they been doing this for…”; “I’d want to know when it was formed, why is formed, where is formed and how is formed.”	15.3%
Collaborative procedures	Procedures governing how the commission conducts inquiry	“What time meeting start, what time they end and budgets. I’d want to know about any funding and would want to know about how many people attend these meetings, the backgrounds of the people within these meetings, how long this meetings go on for, when these meetings occur have these meetings getting bigger…”	26.5%
Procedures of argumentation	Processes of inquiry involving discussion and debate	“…they probably discuss this in depth before publishing it, before writing it up…”; “I would want to know if it was, if everyone was logical when they were speaking or if it was more emotional...”	14.4%
Information flow & access	Manner in which information is shared amongst group of inquirers	“The information would probably be shared amongst the group…. Otherwise you are hiding information from the rest of the commission…”; “…and are there other people allowed to sit in on the meetings and do they involve other, organizations with similar interests to sit in on them.”	1.4%
Publication & communicat-ion	Methods by which the inquiring group publicizes its findings	“I’d want to know what the report, annual report on the safety of genetically modified food, what it says, what the references they use are, if it’s not just the interested members, if there’s actual citations from papers. I’d want to know what these papers are.”	20.5%
Funding	Means by which the inquirers are funded	“…who’s funding it”; “…If it’s being funded I’d want to know if the commission is, a paid position or not.”	6.0%
Conditions on processes involving evidence			
Forensic evidence	Additional forensic evidence (e.g. DNA, fibers and fabrics, video surveillance, etc.)	“Were there any fingerprints or other evidence found at the location?”; “…any other evidence of the sign of break in?”; “Okay, if there is only DNA left the scene…”; “…they should interview people who have actually consumed both types of food and see if they’ve had any adverse reactions.”	38.5%
Generic evidence evaluation	Generic evaluation of evidence	“…the evidence is mostly circumstantial…”; “…I think it’s the most plausible conclusion…”; “…I mean if I was in a jury I would have some reasonable doubt…”; “…that’s still inconclusive.”	47.3%

Other evidence	Other kinds of evidence not mentioned above, or “evidence” in a non-evaluative mode	“…any evidence that can support that he was in the house or not.”; “I would want to know how if the defendant has a limp, he was really able to run 5 miles without anyone catching him”; “…where those people who have those viewpoint got their viewpoints and what they have to back them up with.”	13.5%
		Other	
Other remarks	Requests for further information, personal view or moral judgment, misunderstanding or un-interpretable	“Because you don’t want to wrongly accuse the defendant right, it’s the question---here especially in America, the punishments are really severe in terms of fine and jail time so it would be really wrong to wrongly accuse somebody.”; “I’m against genetically modified food. You should just eat it in its natural form without adding what are they called, growth hormones.”	100%

Appendices

Appendix A: Eyewitness vignette and interview questions

Read the following story and answer the interview questions that follow:

Imagine that you are investigating a criminal lawsuit. The lawsuit involves a plaintiff and a defendant. Your role is to find out whether there is sufficient evidence to convict the defendant. The evidence: The plaintiff has accused the defendant of burglary, specifically the theft of a wallet containing 200 dollars from the plaintiff's home. The plaintiff has stated that he saw the defendant climbing over the back wall of their garden on the night of the crime. The defendant was arrested several days later because he fitted the plaintiff's general description of the culprit. 340 dollars was found at the defendant's home, which is about 5 miles from that of the plaintiff. The plaintiff subsequently identified the defendant in a police lineup, and claims to remember that the person who committed the crime had the same distinctive limp as the defendant. The defendant claims to be innocent, and has testified that he regularly keeps cash in his home.

Your role is to find out whether there is sufficient evidence to convict the defendant.

Eyewitness vignette interview questions

a. If you had to make a decision right now about this case, would you be able to? Why? Why not? 1. Yes. 2. No. 3. Not sure.

b. If you had to decide right now, what decision would you make?

1. The defendant's case is better supported by the evidence.

2. Both cases are equally supported by the evidence (or you cannot tell).

3. The plaintiff's case is better supported by the evidence.

c. How confident are you about this decision?

1. Low confidence. 2. Neither high nor low confidence. 3. High confidence.

d. To find out whether the evidence supports the case of plaintiff or the defendant, what additional information would you need to find out? What would you need to find out? Try and list as many questions you would need to get answered in order to make your decision.

e. What would you need to find out about the plaintiff? Why would you need to find that out? How would this influence your decision? What else would you need to find out about the plaintiff?

f. What would you need to find out about the night of the crime? Why would you need to find out that? How would this influence your decision? What else would you need to find out about the night of the crime?

g. What would you need to find out about the alleged eye-witnessing of the crime? Why would you need to find out that? How would this influence your decision? What else would you need to find out about the eye-witnessing of the crime?

h. What would you need to find out about the plaintiff's general description of the culprit? Why would you need to know that? How would this influence your decision? What else would you need to know about the general description of the culprit?

i. What would you need to find out about the police lineup? Why would you need to know that? How would this influence your decision? What else would you need to know about the police lineup?)

j. What would you need to find out about the defendant? Why would you need to know that? How would this influence your decision? What else would you need to know about the defendant?

k. What else would you need to find out? Why would you need to know that? How would this influence your decision?

Appendix B: Commission vignette and interview questions

Read the following story and answer the interview questions that follow:

The safety of genetically engineered food is currently being debated by the New Jersey Commission on Food Safety (NJCFS). The commission was formed by a group of concerned citizens, all of whom are interested in protecting local food supplies from contamination, as well as protecting the environment from pollutants. The commission aims to achieve this goal by conducting interviews of interested members of the public, who apply to make presentations at the commission meetings. The commission uses the information they gather to publish annual reports on the safety of genetically modified foods. The commission is headed by three commission leaders who run the commission meetings and are responsible for writing up annual reports (the first of which will soon be published).

Your role is to find out whether the findings of the commission should be relied upon.

Commission vignette questions

a. If you had to make a decision right now about this case, would you be able to? Why? Why not? 1. Yes. 2. No. 3. Not sure.

b. If you had to decide whether to accept the findings of the commission right now, what decision would you make?

1. Accept the findings of the commission.

2. Neither accept nor reject the findings of the commission (or you cannot tell).

3. Reject the findings of the commission.

c. How confident are you about this decision?

1. High confidence. 2. Neither high nor low confidence. 3. Low confidence.

d. To find out which whether to accept the findings of the commission, what additional information would you need to find out? What would you need to know? Try and list as many questions you would need to get answered in order to make your decision.

e. What would you need to find out about the members of the commission (who makes it up)? Why would you need to know that? How would this influence your decision? What else would you need to know about the commission?

f. What would you need to find out about how the commission was formed? Why would you need to know that? How would this influence your decision? What else would you need to know about how the commission was formed?

g. What would you need to find out about how the commission is run during its meetings? Why would you need to know that? How would this influence your decision? What else would you need to know about the commission is run?

h. What would you need to find out about how the commission generates its annual reports? Why would you need to know that? How would this influence your decision? What else would you need to know about how the commission generates its annual reports?

i. What would you need to find out about how the commission selects and conducts interviews of people? Why would you need to know that? How would this influence your decision? What else would you need to know about the commission selects and interviews people?

j. What else would you need to find out about the commission…? Why would you need to know that? How would this influence your decision? What else would you need to know about the commission?

Appendix C: Medical and judicial essays and written questions

Doubting hearts debate cholesterol. April 10, 2004. By Jerome Burne. There is almost no connection between the amount of cholesterol in your blood and your risk of a heart attack. Not only that, if you don't already have heart disease, you probably won't live any longer if you bring your cholesterol level down. Finally, statins, the cholesterol-reducing drugs we are all being urged to take, are of little use to women. These are just a few of the highly controversial claims being made by the Commission on Clinical Policies, a group of researchers who all consider themselves "Cholesterol Skeptics," and who are challenging one of the cornerstones of public health policy: the notion that reducing cholesterol saves lives at risk from heart disease. While any doctor will tell you that if your cholesterol level is higher than five (millimoles a litre) you should bring it down, probably by taking one of the statins family, these researchers disagree. According to the Commission on Clinical Policies, however, this will involve not only a massively increased drug bill when many cheaper options are available, but will benefit only those men who already have a heart condition. Can such views, which fly so directly in the face of the medical establishment, have any basis in fact? The commission makes some challenging points. For instance, they say there is little evidence that a longer life results for those millions of people who for years have taken their drugs and endured cholesterol-reducing diets. A number of trials have found that, even though the number of deaths from heart disease does fall when cholesterol is reduced by a range of means among patients in primary care - that is, at family doctor level - there is often an increase in the overall death rate from other causes. Writing in the *British Medical Journal* at the end of last year, Rebecca Warburton, a professor at the University of Victoria in Canada, reviewed studies of statins and concluded: "Statins in primary prevention have not consistently reduced the incidence of myocardial infarction [heart attack] or stroke. Other studies have even found that, over the age of 50, reducing cholesterol increases the death rate." The notion that cholesterol is linked to heart disease goes back to the middle of last century, along with the idea of bringing cholesterol levels down with a low-fat diet to protect the heart. Both of these ideas have been strongly challenged. For example, plenty of studies show that only 50 per cent of people who develop heart problems have high cholesterol, while a study in the *British Medical Journal* in 2001 found no link between changing fat in the diet and heart disease. "At a global level, the link with cholesterol and heart disease is far more tenuous than is generally supposed," says Malcolm Kendrick, a family doctor from Cheshire, in the north-west of England, who is the most active skeptic in Britain. "For instance, in Russia at the moment, heart attack rates are rising dramatically but their cholesterol levels are the reverse of what we see in the US and the UK. They often have high levels of the so-called 'good' HDL cholesterol and low levels of the 'bad' LDL, but they still keel over from heart disease." Even in the West the link is pretty thin, according to Joel Kauffman, a professor at University of the Sciences in Philadelphia. A review he did of statin use last year pointed out that what does correlate with high cholesterol is age, a major factor in heart disease: "When you correct for age, there is almost no correlation between high cholesterol and heart disease." This challenge comes at a time when governments, the medical profession and the pharmaceutical industry are united in their approval of cholesterol-reducing drugs. Recent British Government figures, for instance, show that heart attack deaths are declining, and part of the credit for this is given to statins. The commission also raised new queries about side effects. Statins are generally described as safe and well tolerated. But the same report concluded that although patients on statins had a 1.4 per cent lower rate of heart attacks, this was cancelled out by a 1.8 per cent rate of "serious adverse events associated with the drug," including cancer. That, they say, is almost certainly an underestimate since only two of the trials provided details of any serious side effects. The researchers on the commission said they had asked the drug producers for the missing data but received no reply. To an outsider what is curious about this debate is that both sides are using the same data; much of the disagreement is based on how you interpret it. Perhaps a new commission is needed on this issue – one that gets all the different disagreeing parties into a room to try to work out their disagreements. The cholesterol hypothesis is unlikely to be abandoned in a hurry, given the weight of financial and political muscle behind it. But the commission of skeptics have raised questions that could have an impact on the way we think about heart disease.

Taken from: http://www.smh.com.au/articles/2004/04/09/1081326926651.html?from=storyrhs

Write a short essay (300–500 words) in response to the article focusing on the following questions: 1. What do we know about cholesterol, and how do we know it? 2. What kinds of steps could we take to improve our understanding of the effects of cholesterol?

Ohio Inmate Gary Thompson Executed.

By MICHAEL GRACZYK. HUNTSVILLE, Ohio (AP) — A defiant Gary Thompson struggled as he went to his death, insisting he was innocent moments before an execution that presented Gov. Jim Lancaster with the loudest outcry over capital punishment since he began his run for president. Thompson, 36, received a lethal injection Thursday night for killing a man outside a Houston supermarket in 1981. The state parole board and appeals courts rejected Thompson's arguments that he was convicted on shaky evidence from a single eyewitness and that his trial lawyer did a poor job. Lancaster, said he supported the execution, the 35th during his 5 1/2 years in office. He said the case had been reviewed by 33 judges in 19 years. "Mr. Thompson has had full and fair access to state and federal courts," Lancaster said less than an hour before the execution. "After considering all the facts, I'm confident justice is being done." Gary Thompson promised to "fight like hell" before his death and he did. He resisted coming out of his cell and it took five officers to strap him to a gurney. "This is what happens to a black man — genocide in America," Thompson said in the death chamber, almost spitting out and shouting his words. He called for a moratorium on the death penalty in Ohio, which leads the nation with 222 executions since capital punishment resumed here in 1982. "This is nothing more simple than murder, state-sanctioned murder in America," Thompson barked. "They know I'm innocent. They won't acknowledge it." Last-minute court activity delayed the execution for almost three hours as hundreds of demonstrators ranging from socialists to the Ku Klux Klan waited outside behind barricades and guarded by officers wearing riot gear. There were nine arrests. Thompson was pronounced dead at 8:49 p.m., eight minutes after the lethal drugs began flowing into his arms. Afterward, the Rev. Jesse Jackson, asked by Thompson to witness his death, said: "They are using him as a political sacrifice." Thompson was convicted of killing 53-year-old Bobby Lambert of Tucson, Ariz., in a holdup outside the supermarket. He pleaded guilty to 10 robberies during a weeklong rampage around the same time but said he was innocent of the murder. No physical evidence tied Thompson to the killing, and ballistics tests showed that the gun he had when he was arrested was not the murder weapon. But the lone witness who identified him, Bernadine Skillern, has never wavered. Skillern, who was waiting in her car outside the market while her daughter ran inside, saw the holdup from about 30 feet away. "I don't feel joy and I don't feel sadness," she said after the execution. "I only feel relief. I hope to get back to my privacy, put this incident behind me and now move on." Thompson also argued that his lawyer during the trial, Ron Mock, should have introduced other witnesses who would say he was not the killer. But those witnesses initially told police they couldn't identify the killer, and prosecutors said they were not actual eyewitnesses. Mock has said Thompson gave him no names of alibi witnesses before the trial. The lawyer said Thompson told him only that he had spent the evening with a girlfriend whose description and address he could not remember. The debate over Thompson's case came amid growing questions about the death penalty. Illinois Gov. George Ryan has placed a moratorium on executions. Lancaster could not stop Thompson's execution because the inmate received a one-time, 30-day gubernatorial reprieve from Lancaster's predecessor, Bert Richards. Preceding the 11th-hour legal maneuvers, the Ohio Board of Pardons and Paroles rejected by a 17-3 vote Thompson's request for a 120-day reprieve. The 18-member panel, with one member absent from the vote, also voted 12-5 against a commuted sentence and 17-0 against a pardon. A few hours later, the U.S. Supreme Court rejected Thompson's appeal on a 5-4 vote. Outside the prison, eight people were arrested for breaking through police lines and a juvenile was arrested for assaulting a prison administrator. Other activists burned American flags. Protests were also held in Austin and as far away as San Francisco and Northampton, Mass. "By no way are we happy Gary Thompson is dead," said Lambert's son, Stephen. "He put himself in that situation. We didn't put him there." He rejected the notion that executions offer "closure" to families of murder victims. "It's not over, because my dad's still dead," he said.

Taken from: http://www.bhpioneer.com/article_d898e370-6ac8-5210-9e5f-54a38986898d.html

Write a short essay (300-500 words) in response to the article focusing on the following statement: 1. What do we know about Gary Thompson, and how do we know it?

References

Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95(3). 518-542.

Becker, K. (2009). Reliabilism. *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, http://www.iep.utm.edu/, 25 August, 2013.

Buckland, L.A., & Chinn, C.A. (2015) Epistemology. *Encyclopedia of Science Education*. Gunstone, R. (Ed.) Springer Press.

Chinn, C.A., Buckland L.A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: arguments from philosophy and psychology, *Educational Psychologist*, 46, 141–167.

Chinn, C. A., Rinehart, R. W. & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In Rapp, D. N. & Braasch, J. (Eds.). *Processing Inaccurate Information*.

Descartes, R. (1996). *Descartes: Meditations on first philosophy: With selections from the objections and replies*. Cambridge University Press.

Dretske, F. (2000). *Perception, knowledge, and belief*. Cambridge, England: Cambridge University Press.

Duschl, R.A., & Wright, E. (1989). A case study of high school teachers' decision making models for planning and teaching science. *Journal of Research in Science Teaching*, 26, 467–501.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, 23(6), 121-123.

Goldman, A. I. (1986). *Epistemology and cognition.* Cambridge, MA: Harvard University Press.

Goldman, A. I. (1999). *Knowledge in a social world*. Oxford: Clarendon Press.

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, 43(3), 142–160.

Hennessey, M. N., Murphy, P. K., & Kulikowich, J. M. (2013). Investigating teachers' beliefs about the utility of epistemic practices: a pilot study of a new assessment. *Instructional Science*, *41*(3), 499-519.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, *Review of Educational Research*, 67, 88-140.

Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. NJ: Erlbaum: Mahwah.

Intons-Peterson, M.J. & Best, D. L. (1998). *Memory distortions and their prevention*. Mahwah, NJ: Erlbaum.

Jehng, J. C. J., Johnson, S. D., & Anderson, R. C. (1993). Schooling and students′ epistemological beliefs about learning. *Contemporary educational psychology*, 18(1), 23–35.

Kang, N. H., & Wallace, C. S. (2005). Secondary science teachers' use of laboratory activities: Linking epistemological beliefs, goals, and practices. *Science Education*, 89(1), 140–165.

Kelly, G. J., & Duschl, R. A. (2002). Toward a research agenda for epistemological studies in science education. Paper presented at the Annual Meeting of NARST 2002, New Orleans, LA.

Kitcher, P. (1993). *The advancement of science.* Oxford: Oxford University Press.

Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition and change*. Chicago: University of Chicago Press.

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In Hofer, B. K., & Pintrich, P. R. (Eds), *Personal epistemology: The psychology of beliefs about knowledge and knowing*, Erlbaum, Mahwah, NJ. 121–144.

Lederman, I. L. (1985). Science and technology in Europe: A survey. *Science and Public Policy*, 12, 131–143.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*(4), 361-366.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Edwards, M. N. (2007). Epistemological threads in the fabric of conceptual change research. In S. Vosniadou, A. Baltas, & X. Vamvakoussi (Eds.), *Reframing the conceptual change approach in learning and instruction*. pp. 105–122. Oxford, UK: Elsevier.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Hennessey, M. N. (2012). Examining epistemic frames in conceptual change research: implications for learning and instruction. *Asia Pacific Education Review*, *13*(3), 475–486.

Nozick, R. (1981). *Philosophical explanations*. Harvard University Press.

Nott, M., & Wellington, J. (1996). Probing teachers' views of the nature of science: How should we do it and where should we be looking. *Research in science education in Europe: Current issues and themes*, 283–293.

Nott, M., & Wellington, J. (1998). Eliciting, interpreting and developing teachers' understandings of the nature of science. *Science and Education,* 7, 579–594.

Plato. (2008). *Theaetetus*. 1st World Publishing.

Perry, W. G., Jr. (1968/1999). *Forms of intellectual and ethical development in the college years: A scheme.* San Francisco, CA: Jossey-Bass.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. Science Education, 89, 634–656.

Sandoval, W. A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching,* 40, 369–392.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of educational psychology*, 82(3), 498.

Sieck, W. R., & Mueller, S. T. (2009). Cultural variations in collaborative decision making: driven by beliefs or social norms? In *Proceedings of the 2009 international workshop on Intercultural collaboration* (111–118). ACM.

Simons, D. J., & Chabris, C. F. (2011). *What people believe about how memory works: A representative survey of the U. S. population*. PLoS One, 6, e22757.

Sosa, E. (2001). Reliabilism and intellectual virtue. In H. Kornblith (Ed.), *Epistemology: Internalism and externalism*. (147–162). Malden, MA: Blackwell.

Tobin, K. & McRobbie, C. (1997). Beliefs about the nature of science and the enacted science curriculum. *Science and Education*, 6, 355–371.

Tversky, B., & Marsh, E. J. (2000). Biased retellings of events yield biased memories. *Cognitive psychology*, *40*(1), 1-38.

Waters-Adams, S. (2006). The relationship between understanding of the nature of science and practice: The influence of teachers' beliefs about education, teaching and learning. *International Journal of Science Education*, 28*,* 919–944.

CHAPTER III: EPISTEMIC GROWTH IN MODEL-BASED REASONING

Abstract

Practices of scientific modeling are a valuable resource for facilitating effective science learning through inquiry. Crucially, model-based reasoning brings the justificatory practices of authentic scientific inquiry to the heart of the science classroom, and so engenders a more epistemologically sophisticated understanding of science. However, there has been little research into the impact of sustained modeling on the justificatory practices of science learners. In addition, few studies have investigated change in the high-level criteria that guide these practices. This study explores the model-based reasoning of 24 seventh-grade inquiry learners in the classes of four teachers in two schools over a full school year. It investigates change in participants' justificatory practices and in the epistemic criteria that guide their judgments about model quality. During instruction, a modular life-science curriculum spanning multiple content areas integrated modeling activities into daily learning. Students constructed, evaluated, justified, and successively modified models while engaged in collaborative discourse around rich bodies of evidence and communally negotiated sets of criteria for model evaluation. Data sources comprise learners' written justifications of constructed, peer, group and given models, and their comparative evaluations during model choice. The study reveals the changing justificatory strategies adopted by learners, their use of evidence and elaborated reasoning, and the complex array of epistemic criteria that feature in their judgments about model quality. By elucidating these patterns of change in practices and conceptions of model justification, the study aims to bring to light the reasoning and discursive affordances of sustained model-based inquiry learning.

Keywords: inquiry learning; justification, models; model-based reasoning; epistemic cognition

Introduction

A growing movement in science education maintains that to promote the effective learning of science, students require authentic, inquiry-based, and student-driven investigations of relevant content. There is also growing agreement that achieving effective science inquiry learning requires bringing the discursive and epistemic practices of science and scientists to life in the classroom, particularly practices of collaborative argumentation (Duschl, 2002; Osborne, 2010). One critical way in which the language and reasoning of real science can be instantiated in classroom learning is through the promotion of practices of scientific justification (Erduran & Simon, 2004; McNeill & Krajcik, 2012; Osborne, Sandoval & Cam, 2011). This pedagogical approach typically involves the use of scaffolds in coordinating claims with the reasons and evidence that might bear upon them (Berland & Reiser, 2009; McNeill, 2011). Another productive approach involves the integration of models, modeling, and model-based reasoning into learning (Coll, France & Taylor, 2005; Harrison & Treagust, 2000; Schwarz, Reiser, Davis, Kenyon, Achér, Fortus, Shwartz, Hug & Krajcik, 2009). These two foci, on scientific justification and modeling, are strongly complementary. In particular, the use of modeling in the classroom, including both student-developed and authentic scientific models, is likely to provide an engaging way of learning the justificatory practices of real science. Those working at the confluence of these strands of research have further highlighted the importance of the epistemic criteria that underlie judgments of model quality (Pluta, Chinn, and Duncan, 2011). They have argued that building these criteria into classroom discourse provides learners with the conceptual tools necessary for engaging in the kinds of collaborative argumentation that promotes effective inquiry learning.

Given the value of model-based reasoning for inquiry learning, a considerable body of research has investigated practices of classroom modeling and justification (Lehrer & Schauble, 2006; Passmore, Stewart & Cartier, 2009). However, much of this work involves short-term, cross-sectional studies using a pretest-posttest methodology, with few studies tracing fine-grained patterns of change in learners' reasoning practices over sustained periods of learning (Schwarz, Reiser, Davis, Kenyon, Achér, Fortus, Shwartz, Hug & Krajcik, 2009). Extended, ecologically valid studies of this kind are vital, given that successfully integrating these complex practices into everyday learning is likely to take time, both for teachers and learners (Bransford, Brown & Cocking, 2000). Similarly, few studies have sought to trace change in the criteria that underlie learners' judgments about model quality (Pluta, Chinn, & Duncan, 2011; Schwarz & White, 2005). Studying the meta-knowledge implicit in epistemic criteria is vital for charting the trajectories of conceptual change that different individuals and groups demonstrate during model-based learning. This fine-grained charting of developmental trajectories promises to advance our understanding of the ideal learning progressions for scientific modeling (Schwarz et al. 2009).

The current study thus extends research in this area by investigating patterns of change in learners' justificatory practices and modeling criteria at a fine grain-size and over an extended period of learning. It analyzes the evolving reasoning practices of 24 seventh-grade students, involved in a yearlong, criteria-infused and model-based inquiry-learning environment. The study examines participants' written justifications of life science models, with the aim of exposing the epistemic criteria underlying their judgments about model quality. It seeks to trace detailed patterns of change in their modeling conceptions and skills of scientific justification. Achieving this aim requires documenting the ways in which modeling criteria feature in their practices of justification over an extended period of inquiry learning. For example, some learners might regard a model's provision of mechanism as the defining virtue of good models, and so prefer models that feature mechanisms over those with a far better fit with the evidence but without a detailed mechanism. The justifications learners give for their judgments about model quality will likely reveal their conceptual grasp of models and modeling; concepts that are likely to change over time as they develop modeling expertise.

This study continues the investigation of middle-school science inquiry learning in Pluta, Buckland, Chinn, Duschl, and Duncan (2008), Buckland and Chinn (2010), and Pluta, Chinn, and Duncan (2011). In particular, Pluta et al. (2011) identified the range of epistemic criteria used by middle-school science students as they explicitly reflected on model quality. The study advances the field by considering concurrent (and possibly inter-dependent) change in three vital areas of model-based cognition – epistemic criteria for model quality, use of evidence in justifications, and skills of elaborated reasoning.

The central research questions of the study are thus:

1. What higher-level criteria for model quality feature in participants' justifications, and how do these change over the course of a year?
2. How do participants use evidence in their model justifications, and how does this change over the year?
3. What degree of elaboration do students' model justifications exhibit, and how do their elaborative practices change over the year? That is, how much detail do they provide in explicating the link between their model and the relevant evidence or reasons provided?

By investigating change in these three interlocking aspects of participants' model justifications this study aims to present a more detailed and holistic picture of the ways in which justificatory practices develop over the course of sustained model-based inquiry learning. In the remainder of the introduction, we briefly explore each of the distinct strands of research that this study brings together: practices of scientific justification, scientific modeling in the classroom, and the focus on epistemic criteria for model quality.

Practices of Justification in Science Learning

The justification of claims, explanations, and models is a critical scientific activity, one that should be ubiquitous in science learning through inquiry. While scientific explanation involves the development of integrative, unifying descriptions of how things work that cohere with the evidence, justifications are inherently evaluative, and involve the provision of reasons in support of those explanations (Nagel, 1961). The ability to recognize, construct, and evaluate justifications thus represents a crucial learning outcome. This is particularly the case for pedagogies that recognize science as both interconnected disciplinary content knowledge and as a set of reasoning practices and epistemic values (Schweingruber, Keller & Quinn, 2012).

Common to all acts of justification is the provision of reasons in support of conclusions. Given that justifying one's beliefs involves giving reasons in their support, justification is a critical feature of scientific argumentation more generally (Kuhn, 1991). The *Next Generation Science Standards* (NGSS) recommends that science learners practice "engaging in argument from evidence" which requires them to "construct, use, and/or present an oral and written argument or counter-arguments based on data and evidence" (National Science Teacher Association, 2013, p. 6). These vital practices of argumentation necessarily involve skills of justification.

Practices of justification are also particularly important for successful learning through inquiry, given the importance of argumentative discourse amongst collaborative inquirers. Osborne (2010, 2014) has argued that promoting practices of justification is essential in science-learning environments because understanding science requires both knowledge of content and the ability to see how it is that we accept those claims to knowledge. Osborne has claimed that: "Any science education that offers students only a conceptual understanding of science without explaining how we know what we know or why we believe what we do leaves students without any knowledge for the epistemic basis of belief" (2014, p. 580). Skills of justification are what allow learners to make sense of how scientists know what they claim to know.

Besides its importance for argumentation and learning through inquiry, justification is also a fundamental dimension of epistemic cognition. Epistemic cognition is a form of metacognition targeted at epistemic ends like knowledge and understanding. It involves beliefs about knowledge, knowing, and knowledge-generating practices and institutions (Buckland & Chinn, 2015). Research in this area investigates the ways in which people conceptualize and use epistemological notions, particularly justification. It aims to trace the role of tacit and explicit epistemic beliefs in reasoning and decision-making. Perhaps the most widely used framework for studying epistemic cognition, that of Hofer and Pintrich (1997, 2002), features justification as one of four core dimensions, alongside simplicity, certainty, and sources. This framework individuates learners' justificatory strategies in terms of their reference to authority, personal experience, and rules of inquiry. Many others, like Greene, Azevedo, and Torney-Purta (2008) and Muis (2008), have also recognized the centrality of justification to epistemic cognition.

Chinn, Buckland and Samarapungavan (2011) have advocated for an expanded framework for developing models and measures of epistemic cognition (see Introduction and Chapter 2). In particular, they have advocated for the implementation of more fine-grained and integrated analyses of justificatory practices, with a particular focus on situated and ecologically valid justifications. The current study aims to implement these recommendations by targeting highly contextualized and authentic episodes of justification, firmly situated in middle-school inquiry learning environments. It realizes a fine-grained level of detail by gathering multiple instances of justification for each participant, analyzing these using a rich analytic framework, and considering skills and strategies of justification (e.g. ability to develop well-elaborated justifications) in conjunction with conceptual understanding (e.g. criteria for model quality).

Research has shown that learners can exhibit various strategies of justification, including reference to data and empirical evidence, presentation of scientific norms and concepts, appeal to authority and/or personal experience, and depiction of mechanism (McNeill & Krajcik, 2012; Osborne et al., 2004; Sandoval & Cam, 2011). It has also identified a range of the strengths and weaknesses in learners' capabilities. At a very general level, research has shown that student discourse typically features a great many unjustified claims (Jiménez-Aleixandre, Rodríguez & Duschl, 2000). Similarly, researchers have shown that students do not tend to spontaneously adopt skills of scientific justification, but require much practice and explicit instruction (Osborne, Erduran & Simon, 2004). Other deficits that have been revealed include a difficulty in consistently differentiating evidence and theory (Kuhn, 1991), an over-reliance on personal judgment over evidence (Hogan & Maglienti, 2001), and difficulties in formulating justifications and identifying evidence (Sadler, 2004). Sandoval (2003) and Sandoval and Millwood (2005) found that even when student were able to consistently identify evidence, they battled to select the evidence appropriate to or adequate for the justification they proposed. Bell and Linn (2000) further found that even when learners successfully developed justifications that invoked the appropriate evidence, they had great difficulty in saying why they chose the evidence they did. Similarly, McNeill, Lizotte, Harris, Scott, Krajcik, and Marx (2003) found that middle-school students did not generally succeed in developing quality links between their claims and the evidence they identified. For example, they were generally unable to articulate the scientific principles or rules of inference that those links depended upon.

Given these deficits, the field has investigated ways in which to promote the development of learners' skills of justification. One very popular approach is the *claim*, *evidence*, and *reasoning* instructional scaffold (Berland & Reiser, 2009; McNeill, 2011; Songer & Gotwals, 2012). Originally developed from Toulmin's (1958, 2006) scaffold for argumentation, this framework was intended to support learners' development of scientific explanation through the provision of evidence and reason for claims. On this scaffold, the *claim* is a propositional claim intended to answer an inquiry question. *Evidence* is the data used in support of the claim. *Reasoning* involves an explication of the link between the *claim* and *evidence* using scientific ideas and concepts, as well as showing how the data identified counts as supporting evidence. This relatively structured approach to promoting argumentation contrasts with approaches without such explicit scaffolding (e.g. Roseberry, Oganowski, DiSchino, and Warren, 2010).

These researchers have described the claim-evidence-reasoning scaffold as a tool for constructing scientific explanations. However, given that linking evidence to claims through reasoning is primarily a *justificatory* rather than explanatory practice, this approach has come under criticism. Osborne and Patterson (2011) have argued that the scaffold fails to properly distinguish explanation and argumentation, stating, "Lacking a well-defined intellectual construct students are in danger of confusing the goals of argument and explanation, omitting vital elements of both" (p. 636). This is of particular importance given that Toulmin's framework provides the conceptual basis for the claim-evidence-reasoning scaffold, yet it is a tool for structuring productive argumentation rather than explanation. Berland and McNeill (2012) have broadly agreed with this criticism, stating, "argumentation and explanation are distinct scientific practices that are often treated as one in science education" (p. 811). Nonetheless, they have remained committed to the scaffold as originally construed, and have argued that this promotes consistency with widely accepted science standards and better aligns with school culture (McNeill, Lizotte, Krajcik and Marx, 2006).

The current study regards explanation and justification as involving crucially different norms and practices. To explain is to provide unifying, often causal accounts of how and why things happen (e.g. using models) – this is a fundamentally *descriptive* project. In contrast, to justify is to give reasons and evidence in support of claims and beliefs – an essentially *evaluative* project. Science clearly involves both explanatory and justificatory practices, with arguments in support of particular explanations often interwoven with those explanations. However, it is nonetheless crucial to encourage science learners to practice both the descriptive and evaluative components of scientific reasoning, and to develop the strategies and expertise specific to each. This study thus aims to extend research on the scaffolded integration of practices of scientific justification in inquiry learning, while avoiding the conflation of explanatory with justificatory modes of reasoning.

Berland and McNeill (2010) have developed a learning progression for argumentation with relevance to the current study. This progression focused on three inter-related elements of argumentation: the instructional context, the argumentative product, and the argumentative process. For Berland and McNeill, the primary value of argumentation for learning is its engagement of learners in the construction and justification of knowledge claims. They suggest that, “developing a classroom culture and norms is also essential for supporting student engagement in the argumentative process” (2010, p. 789). To advance learners’ justificatory skills, these norms should facilitate students’ attention to the questioning, evaluation and revision of the ideas of their peers. Their proposed progression thus depends crucially on learners’ understanding these norms of argumentation, their recognition of when the learning environment demands their participation in argumentation, and the complexity and support of instructional materials.

By focusing on instructional context and, argumentative products and processes, the current study aligns with Berland and McNeill’s (2010) understanding of how skills of justification progress. The instructional approach integrated epistemic criteria for model quality into instructional scaffolds and daily inquiry-driven discourse. These criteria thus served as the norms of scientific argumentation adopted into the classroom culture, a culture involving constant collaborative evaluation of peers’ models, model revisions and model justifications.

Model-Based Inquiry Learning of Science

Scientific models serve to explain how and why things happen, and are subject to successive revision in light of experiment, evidence, and critical debate (Kitcher, 1993; Giere, 1988, 2004;

Godfrey-Smith, 2006; Nersessian, 2002). There is wide consensus that modeling is an important part of authentic scientific practice, and that model-based reasoning provides an effective way of learning science through inquiry (Duschl, 2002). The current policy context reflects this consensus, with US and international standards for science learning recommending a model-based pedagogy. For example, the NGSS identified eight essential practices of science and engineering for learners, including “developing and using models,” described as the ability to “develop, revise, and/or use a model based on evidence to illustrate and/or predict the relationships between systems or between components of a system” (National Science Teacher Association, 2013, p. 6).

Researchers have investigated practices of classroom modeling to determine their role in promoting the effective learning of science. For example, Clement (2000) has suggested that models, when internalized by science students, generate psychologically satisfying explanations and embody a flexible form of knowledge that encourages transfer between problems. Gobert and Buckley (2000) and Buckley (2000) have claimed that models act as “central guideposts” in promoting student-directed inquiry, and so present a point around which new information can be organized and new questions generated. They have suggested that models help to sustain a virtuous sequence of model construction, critique, and revision. Similarly, Gobert (2000) has linked model-based reasoning and conceptual change at the fifth-grade level, with participants showing improved conceptions of complex causal sequences. Lehrer and Schauble (2006) have maintained that the public nature of models as inscriptions helps to make student thinking accessible to teachers, which promotes the continuous informal assessment of learning.

Some researchers have argued that model-based instruction promotes understanding of science content. This includes discipline-specific research, for example, the investigation of physics understanding of White and Frederiksen (1998). It also includes topic-specific work, for example, evolutionary modeling of Chinn and Buckland (2011). Still others have argued for the domain-general applicability of modeling practices (Krajcik, McNeill, & Reiser, 2008). A group of researchers has investigated the *epistemological* underpinnings of students’ use and understanding of models. For example, Harrison and Treagust (2000) have claimed that secondary students tend to adopt a very simplistic view of the relationship between models and the phenomena they represent. On this view, models bear a one-to-one correspondence to the aspects of the world that they represent, and addressing these views requires epistemologically informed instruction. Coll, France, and Taylor (2005) have argued that attaining modeling fluency requires the development of a dual appreciation of the epistemic basis of scientific models. This dual appreciation includes an understanding of models both as representations of reality and as productive tools for reasoning. Research in this field has thus shown a variety of ways to leverage model-based reasoning to advance important learning goals (see also Grosslight, Unger, Jay & Smith, 1991; Gutwill, Frederiksen & White, 1999; Raghavan & Glaser, 1995; Sandoval & Reiser, 2004; Schwarz et al., 2009; Toth, Suthers & Lesgold, 2002; Treagust, Chittleborough & Mamiala, 2002; White, 1993; White & Frederiksen, 1998; and Windschitl, Thompson & Braaten, 2008).

Schwarz et al. (2009) have developed a learning progression for scientific modeling with a focus on both the practices of modeling and on the meta-knowledge of the nature and purpose of models. Their learning progression bears similarities to the account of Coll, France, and Taylor (2005) in its dual focus on the role of models as predictive and explanatory tools as well as models as revisable in light of improved understanding. Schwarz et al. found that as their understanding of models progresses, learners shift from a conception of models as primarily illustrative and descriptive, to sophisticated views of the explanatory function of models. These conceptions encompass models as tools for sense making and knowledge production. The current study extends this work by drilling deeper into the modeling meta-knowledge of learners. In particular, the criteria that function as norms of classroom culture developed through extended and repeated discussions about how the nature and purpose of models determines criteria for model quality.

In spite of promise of modeling for promoting effective science learning, few studies have specifically investigated change in learners' justificatory practices that are associated with model-based learning. One research strand that has interwoven the emphasis on models with that of justification has done so by focusing on learners' epistemic criteria for model quality.

Epistemic Criteria

Epistemic criteria for models are the central concepts used in the evaluation of models. These evaluative concepts provide the basis for judgments of how good models are (e.g., how useful, accurate, or explanatory) and thus for model justification. For practicing scientists, they provide a vocabulary for the critique, selection, and refinement of competing models. Epistemic criteria for models include evaluations of models as good because they are simple, clear, coherent, parsimonious, and consistent with the evidence; they make novel and testable predictions, cohere with the larger body of theoretical knowledge and have significant explanatory power. Modeling criteria are a subset of the more general norms and standards used to evaluate claims to scientific knowledge (Kuhn, 1977). Other criteria that have been studied include criteria for good evidence, theory-choice and scientific conclusions (e.g. Hogan & Maglienti, 2001; Penner, Giles, Lehrer, & Schauble, 1997; Samarapungavan, 1992; Smith, Maclin, Houghton, & Hennessey, 2000; Penner, Giles, Lehrer, & Schauble, 1997; and Schwarz & White, 2005).

Epistemic criteria thus embody key norms of scientific cognition, and provide an effective way of integrating these norms into learning through collaborative argumentation (Bricker & Bell, 2005; Duschl, 2008; Duschl & Osborne, 2002; Duschl, Schweingruber, & Shouse, 2007; Kelly, 2008; Lehrer & Schauble, 2006). Criteria are important for understanding how scientific norms attain metacognitive status among learners, and are thus adopted as tools for the collaborative regulation of discourse. Criteria-infused discourse thus provides a rich, natural context for students to employ skills of scientific justification. Given that learner's model justifications serve to express their underlying conceptions of model quality, these justifications are an excellent place to investigate the epistemic criteria that guides their model-based reasoning.

In previous work, we have investigated model-based inquiry with a focus on epistemic criteria (Buckland & Chinn, 2010; Pluta, Buckland, Chinn, Duschl & Duncan, 2008; Pluta, Chinn & Duncan, 2011). In particular, Pluta et al. (2011) sought to identify the modeling criteria adopted by 324 middle-school students engaged in model-based inquiry learning. These learners were able to articulate a complex array of criteria in response to explicit questioning. For example, they evaluated models in terms of their structure, clarity, explanatory power, use of evidence and accuracy. However, this study focused only on participants' reflective judgments about criteria for model quality, and it elicited these judgments before the learners' had engaged in model-based, criteria-infused instruction. The current study thus advances the study of modeling criteria begun in Pluta et al. (2011) in several ways. First, it investigates criteria over an extended period of model-based learning, rather than at one or two points in time. This reveals detailed information about how these conceptions might have changed over time as learners' gained modeling proficiency. Second, it studies the role of criteria in actual classroom learning, rather than in learner's reflective judgments. It is thus sensitive to learners' tacit beliefs about model quality that might not be subject to reflective articulation, and thus better ecological validity. Third, it investigates interlocking features of justificatory expertise, including conceptual insights into criteria, use of evidence and ability to develop well-elaborated justifications.

The current study therefore promises to advance understanding of how middle-school learners' underlying conceptions of model quality change during the course of model-based learning, and to provide teachers with new ways of promoting effective collaborative inquiry.

Method

Context of the Study

The Promoting Reasoning And Conceptual Change in Science (PRACCIS) research project provides the larger context of this study. This was a National Science Foundation funded, multi-year microgenetic study of seventh-grade students learning to reason with and about scientific models and evidence (see Chinn, Pluta, Buckland, Rogat, Difranco & Witham, 2010). Although the PRACCIS project predated the NGSS, the goals and practices promoted by the project align well with these standards, given their focus on models and argumentation.

The PRACCIS project aimed to advance theoretical conceptions of how students learn to reason in the complex learning environments of science classrooms. It developed middle-school science instruction integrating model-based reasoning into daily life science learning, and used microgenetic methods to investigate its impact. PRACCIS instruction included extended reasoning seminars, engaging students in collaborative argumentation about scientific models and evidence. These seminars featured student-driven, collaborative inquiry using life-science modeling alongside rich bodies of evidence. The reasoning strategies promoted include the building, evaluating, revising, comparing, and justifying of models, as well as their coordination with (sometimes conflicting) evidence. Participating learners also designed and evaluated experiments, made and tested predictions and explanations, and considered socio-scientific problems involving measurement, estimation, and sample size. Participating teachers engaged in professional development to advance their understanding of effective model-based inquiry.

A crucial component of PRACCIS instruction was to make the epistemic criteria for model quality explicit for the whole classroom. Recurring whole-class discussions elicited a public, collaboratively developed, defeasible and so successively modified set of modeling criteria (e.g. Appendix B). These discussions explored the use, purpose, and value of scientific modeling. The resulting criteria provided a framework that guided model-based reasoning in daily learning, including model evaluations, model comparisons, and the justification of these judgments. Teachers were encouraged to allow students a substantial role in the co-construction and refinement of the class criteria, to encourage their ownership and adoption. Examples of learners' model quality criteria include "to make the information simpler," "must be 'real' – have evidence to support it," and "can be built off of /worked on - adjustable."

Participants

In the year of this study, the PRACCIS project included approximately 350 focus participants in the microgenetic study, with another approximately 350 receiving the instructional treatment (including pretests and posttests of reasoning). These participants attended the classes of seven teachers in five New Jersey school districts. The current study selected 24 microgenetic participants from eight classrooms of four teachers in two schools. Situated in two very different school districts of suburban New Jersey, the two schools typically attained starkly different levels of success on statewide tests of math and reading proficiency. They also had significantly different socioeconomic and racial compositions (Pluta et al., 2011). In School 1, with 1% of students on free or reduced-price lunch, 97% of students were white and 90% attained proficiency levels on state tests. In School 2, nearly a third of students qualified for free or reduced lunch, 47% were black and 15% Hispanic, with 65% to 70% reaching proficiency on state tests.

Table 11 shows the distribution of the 24 students across teachers and schools, as well as the kinds and order of instructional units they completed. Two central criteria featured in the selection of participants. First, the selection process drew balanced numbers of male and female learners. Second, performance on a model-based reasoning post-test ensured the final participant pool featured equal numbers of below-average and above-average participants. The process of selection excluded students with large amounts of missing data, either due to nonattendance or lost data.

Materials

The data sources for this study include:

- Written justifications of student-generated, peer and researcher-developed models and model revisions over the course of the year.
- Model-evidence diagrams embedded within learning materials and assessments.
- Class-constructed and revised lists of modeling criteria.

The curricular content of the study included learning modules on photosynthesis, cellular respiration, cell membranes, mitosis, genetics and the cardiovascular system. Each featured multiple models developed by students, teachers, and researchers, and addressed the central life-science phenomenon under study. Associated with these models was a range of evidence; this was in the form of realia, data tables and graphs, class demonstrations, experiments, and written and video descriptions of studies and their findings. Video, small-group audio and written individual and group worksheets captured participants' engagement with these materials.

The model-evidence diagram scaffold, described in Buckland and Chinn (2010), provided another source of model justifications. To complete a model-evidence diagram, students drew a set of arrows between models and sets of evidence, with the arrows indicating relations of evidential support, contradiction, and relevance. Students were then required to defend the arrows they selected by providing justifications. Appendix A shows a completed model-evidence diagram used as one of many assessments of reasoning embedded within the modules.

A diverse range of prompts served to elicit justifications from participants. The motivation for this diversity was to encourage earners to recognize requests for model evaluation presented in a wide variety of formats, as well as to avoid monotony in the use of just one or two prompts. Prompts included:

- "Give reasons for your model."
- "Which model do you think is best? Explain what makes it the best."
- "Is this a good model? Why or why not?"
- "Explain why your model might be better than other models."
- "Provide as much evidence as you can for why this model is the best one."
- "Which of the three models we have just decided on do you think is best? Why do you think it is best?"

Table 12 lists and describes the set of instructional units analyzed and lists the justifications included in the analysis. Figure 19 through 21 shows an extract from the Cellular Respiration unit as an example of study materials. This includes an example of a student-constructed model and associated prompt and justification (Figure 19), a set of given models from the same unit (Figure 20) and several pieces of evidence (Figure 21). The core materials of the study comprise approximately 20 written model justifications per student drawn from this dataset.

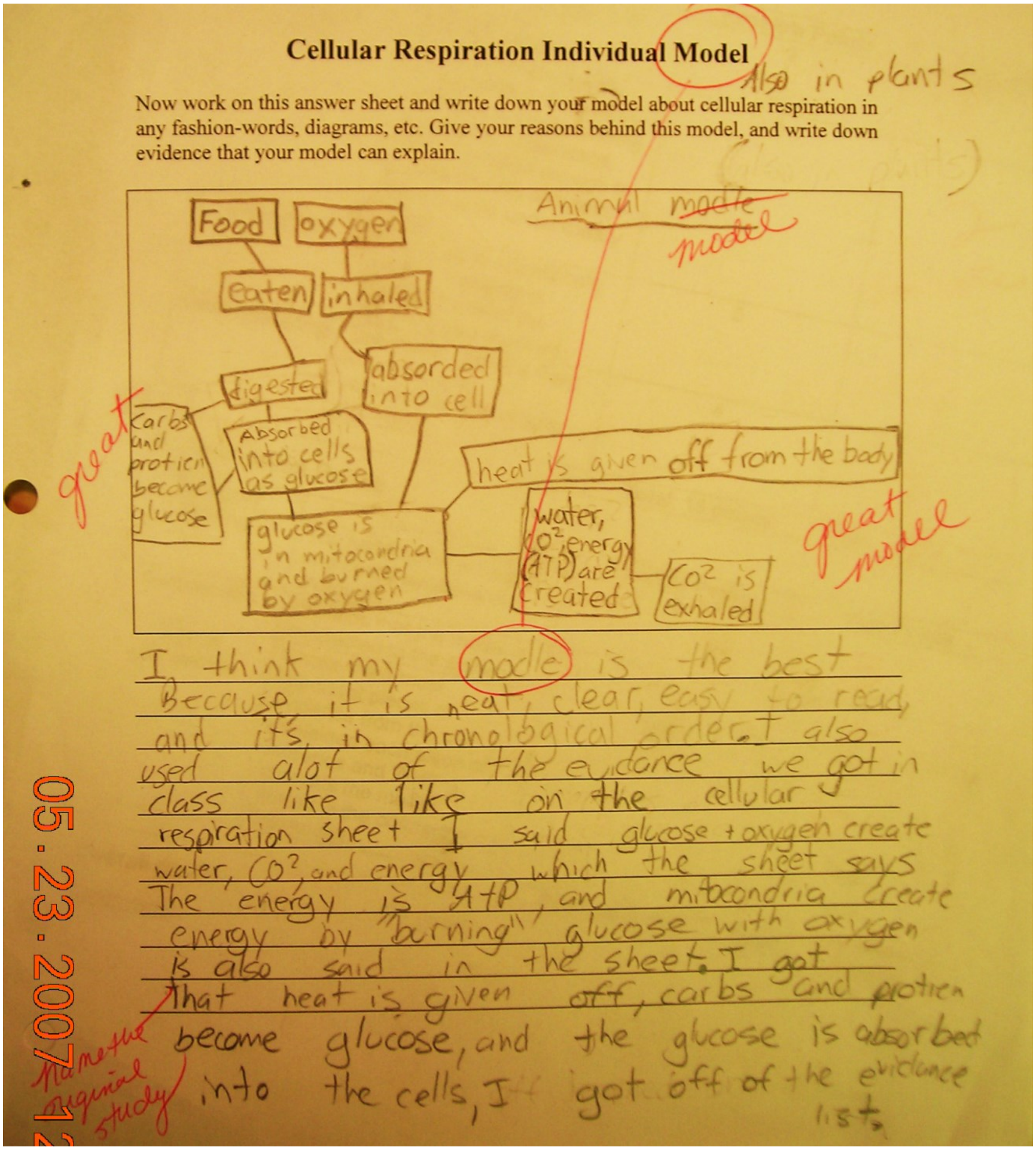


Figure 19. A student model and justification. This figure shows a model, prompt and associated justification from the Cellular Respiration unit.

Examine these three models; then answer the questions at the bottom of the page.

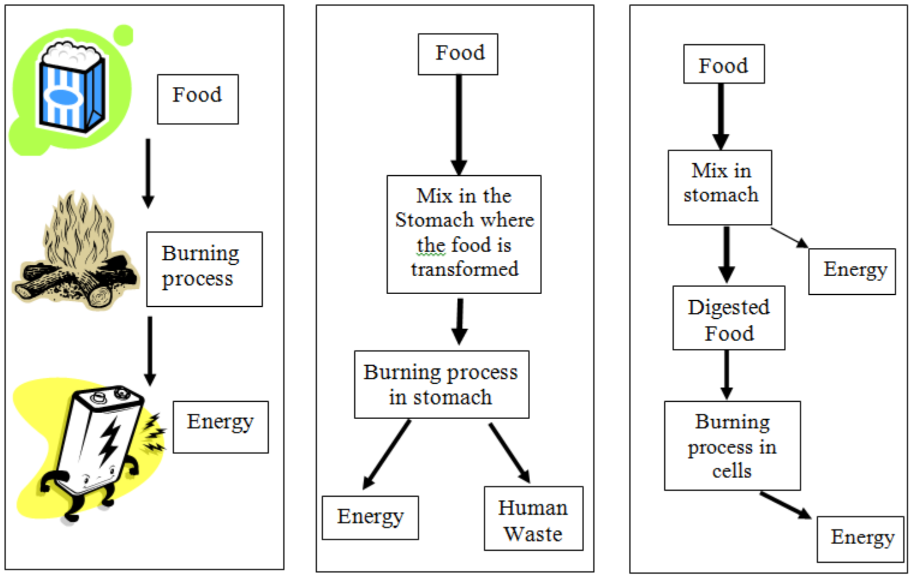
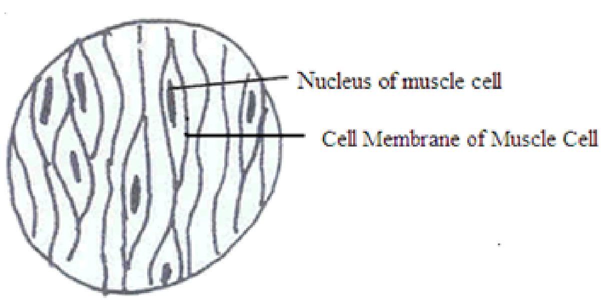


Figure 20. A set of given models. This figure shows a set of models used for purposes of model comparison from the Cellular Respiration unit.

This study describes muscle contraction in the presence of varying levels of sugar and oxygen.

Figure (1) Muscle Cell Contraction during Exercise

Microscopic view of muscle cells



Note: when we use muscles, like doing exercise, the muscle cells in the muscle contract (tighten). The harder the exercise, the faster the muscles contract.

Rate of Muscle Cell Contractions

Amount of Oxygen inhaled		Amount of Sugar (Glucose) eaten		Rate of Muscle Cell Contraction
Low	+	Low	=	Low
High	+	Low	=	Medium
Low	+	High	=	Medium
High	+	High	=	High

Figure 21. Cellular Respiration unit evidence. This figure shows some of the evidence presented to students taken from the Cellular Respiration unit.

Variations in the order and kind of units that teachers implemented complicated the analysis of the resulting dataset. For instance, teacher 2 implemented the Mitosis unit at a different time to the other teachers and did not implement the Cardiovascular systems unit, while teachers 3 and 4 did not implement the Genetics unit. To address this variation, the analysis included five instructional units for each teacher, with participants' justificatory abilities assessed in terms of the relative order of units rather in terms of the particular units themselves. The unit of analysis is therefore the order in which participants completed content modules, and the study traces change in participants' model justifications from early units to later units. Table 11 presents the order of units followed by each teacher.

Coding Scheme

The analysis of participants' justifications involved investigation of their general strategies of justification, the degree to which they elaborated their justifications, and the modeling criteria they adopted. Tables 13 through 15 present the three coding schemes developed. Each table includes a label, a brief explanation, a representative example of each of the component codes and a percentage indicating the total percentage of justifications assigned that code.

Justificatory strategies (Table 13), the first category of codes, refer to the very general ways in which participants responded to justification prompts. First, a set of codes identified responses that failed to provide a genuine justification. The *Withhold judgment* code identified responses that explicitly resisted making justificatory claims. These responses typically denied that there was sufficient information available to make a judgment. The *Re-description* code identified responses that simply re-described the relevant model, rather than evaluating it. The *Explanation* code identified responses that tried to explain rather than justify the model, typically by invoking new ideas that did not feature in the original. The *Non-justificatory* code identified pseudo-justifications, like "The model is best because I said so." The *About* code identified responses that merely asserted that the model and the evidence were about the same thing, frequently using the word "about." Finally, the *Blank* code identified participants' failure to provide a response, and the *Not interpretable* code identified illegible or incoherent responses.

Second, a set of codes distinguished among responses that were at least minimally justificatory. The *Reasons and criteria* code identified justifications that invoked modeling criteria (e.g. simplicity, neatness, explanatory power). The *Evidence* code identified justifications invoking evidence, and captured the number of distinct pieces of evidence mentioned. The *Reference to evidence* code identified justifications that merely included the term “evidence,” without specifying what that evidence consists of (e.g., “The evidence from the studies that we read supports the model.”) Finally, the *Model comparison* code identified justifications that compared competing models. These typically justified models in light of the deficits suffered by the alternatives. As students’ justifications could mention multiple kinds of reasons, these code categories were not mutually exclusive, and a single justification might receive multiple codes.

Elaboration of justification (Table 14) constitutes the second category of codes, and focus on the degree of elaboration of participants’ justifications. Highly elaborated responses typically developed a detailed account of the relationship between a model and the reasons presented in its support. In contrast, minimally elaborated justifications claimed merely that particular reasons provide support, without showing how they do so. This dimension of participant responses thus distinguishes rudimentary justifications from those that are more sophisticated and thoughtful. This code category thus tracked participants’ ability to develop well-elaborated links that make clear why they cited the reasons and evidence they did. First, the *Not applicable* code identified the non-justificatory responses from the first category of codes. The *No elaboration* code identified minimally justificatory responses that that made no effort to show how the reasons cited relate to the model justified. The *Minimal elaboration code* identified responses that included some account, however brief, of how the reasons mentioned relate to the model. Finally, *High elaboration* identified responses that presented detailed, well-developed descriptions of the relation between the reasons (e.g. evidence and/or criteria) cited and the model justified.

Categories of modeling criteria (Table 15) constitutes the final set of codes. These codes identified the epistemic criteria underpinning participants’ model justifications. In developing these codes, the publically posted classroom criteria were first examined, and then broad patterns of similarity served to group these and new kinds of criteria. The *Generic and meta-criteria* code identified justifications that simply claim that a model is "good," "better," "the best" or "meets the criteria." These very vague and general descriptors failed to identify the specific characteristics of a model responsible for its quality – they did not state *why* the model was good. The *Communicative* code identified justifications that focused on the effectiveness of a model in communicating the ideas it embodies. Communicative justifications typically focused on structure (e.g. being simple, having a title, labels and a key), descriptive power (e.g. detailed, informative, clear, comprehensive and understandable), and coherence (e.g. readable, relevant, on-topic and self-explanatory). *Aesthetic* justifications focused on the design and artistic features of a model, i.e., their neatness, beauty, elegance, color, and inclusion of pictures (e.g., "I think model 4. Model 4 has animation pictures. I picked model 4 because people don't always like to just read. People like pictures with captions on it."). *Evidential* justifications involved claims about the empirical character of a model, and included descriptions of models as supported by evidence, studies, facts, background information, or knowledge (e.g. , "Model 3 had the most evidence to support it."). The *Veridical* code identified responses that referred to the veracity of a model as true, accurate, correct, right, proven, and free of falsity (e.g., "It is a proven fact that has been tested."). The *Explanatory* code identified justifications that described models as good in virtue of their explanatory power, claiming that a model reveals how or why particular phenomena occur (e.g., "Evidence #2 supports it because it has a chart and the chart tells you how CO_2 gets in the plant."). *Logic and reason* identified references to the logic or plausibility of models, describing them as rule-governed, sensible, or reasonable (e.g., "It seemed to make more sense."). *Collaborative and effortful* identified justifications that referred to models as the product of the participation and hard work of multiple members of a group (e.g., "I think our model is a good one because our whole group helped make it.") Finally, *No criteria* identified responses that made no mention of any identifiable criteria.

Some coded examples reveal the range of participants' responses. In response to the prompt, "Which model do you think is best? Why?" one student simply stated: "It [model 1] is the neatest." This response did deliver minimally justificatory reasons in support of the model. However, the very general reasons expressed did not make clear how the model embodied the characteristics identified in the justification. This justification thus received the following codes across the three coding schemes: 1. *Reasons and criteria*; 2. *No elaboration*; 3. *Aesthetic*.

In response to the prompt, “Explain why your model is better than other models,” a student stated, “I think this is right because we have evidence of the cells splitting and we know living things don’t function without nutrients. I think the inflatable model is out because no more new cells are being created.” This student engaged in a sophisticated act of justification. Rather than simply listing the valuable features of the model, they mention specific evidence, explicitly link this to the model, and then reject a competing model in light of further evidence. This justification received the following codes: 1. *Reasons and criteria*, *Evidence (1 piece)*, *model comparison*; 2. *Minimal elaboration*; 3. *Veridical*, *Evidential*.

In response to the prompt, “Give your reasons for this model,” a student wrote,

> “My model is good because I used scientific evidence to support it … evidence that the mitochondria make the energy for the cell from our cellular respiration notes. This is important for the model because it is the part of the cell that makes the carbs [sic] into energy. It is where a part of cellular respiration happens. From Figures 4 and 5 … I used the evidence that the more mitochondria and energy makes the person can run faster … It supports my model because it proves that the person with more mitochondria would be ahead. Also … I used the evidence that the cell “burns” glucose with oxygen and produces CO2, energy and water. My arrows from the cell show that they need these to perform cellular respiration. This evidence supports it because it shows that is correct with what the cell needs and produces.”

This extended and highly detailed justificatory response presents multiple pieces of evidence and links each to the model under consideration. The student also provides a high level of elaboration, detailing how the evidence presented provides support. This justification received the following codes: 1. *Reasons and criteria*, *Evidence (3 or more pieces)*; 2: *High elaboration*; 3. *Evidential*, *Veridical*, *Explanatory*. These three examples demonstrate the considerable variability in students’ justificatory competence across the three coding schemes.

Inter-coder Reliability

Coders assigned the three categories of codes to 472 responses to the request for justification. The 24 participants thus generated an average of 19.7 responses, with a mode of 17, a median of 20 and a ranged from 12 to 25 responses. In total, 1761 codes identified patterns of participant response, with an average of 3.73 codes assigned per response. A subset of this dataset served for purposes of calculating levels of inter-rater reliability. First, all data was coded by an initial coder. Then, while working closely with the first coder, a second coder assigned codes to a practice dataset taken from randomly selected participants. Once the second coder demonstrated competence with the coding scheme, they independently assigned codes to 20% of the data corpus. This selection included randomly selected participants yet included students from each teacher. On the three coding schemes, agreement levels were 82%, 78%, and 87% respectively. The first and second coder achieved overall agreement levels of 82% on these assigned codes, indicating a moderate level of inter-rater reliability.

Results

For each of the five ordered instructional units, a percentage identified the number of codes assigned in proportion to the number of opportunities learners had to develop a justification. These percentages apply within each of the three coding categories. General trends were then determined, though statistical tests of their significance were not calculated due to the low number of participants. Table 13 through 15 present the findings of the overall analysis, and figures 22 through 26 present the analysis of change over time.

The Overall Analysis

The overall finding summarized in table 13 shows that 47% of justifications included at least some reference to evidence, and 30% expressed reasons and criteria. For evidence-focused justifications, 36% mentioned one piece, 6% mentioned two pieces, and just under 5% mentioned three or more pieces. Responses that made generic reference to “evidence” (without specifying that evidence) comprised 7% of responses, and simply claimed that evidence was “about” the model comprised 7%. Responses that merely described or explained the model comprised 12% and 21% respectively. Considered overall, 29% of responses provided some reasons, evidence or criteria, yet failed to elaborate how these factors link to the model being justified. Another 34% developed minimally elaborated justifications, with 4.4% developing justifications with highly elaborated links. Of the most significant kinds of modeling criteria mentioned overall, 13% were *Communicative*, 11% were *Evidential*, 16% were *Explanatory*, 4% were *Veridical*, and 60% mentioned no criteria

Evolving Justificatory Practices

Figure 22 shows how participants' reliance on the two main non-justificatory strategies (re-descriptions or explanations of a model) declined considerably over the course of the year. Initially, around 48% of responses involved the non-justificatory reiteration or elucidation of the relevant model, yet this had dropped to 17% of responses by the end of the year. This trend is positive and encouraging.

Figure 23 shows participants' changing pattern of providing reasons and criteria for their models, which was roughly stable over the course of the year, between 20% and 30% of responses. In contrast, Figure 24 shows that participants' use of evidence underwent clear positive change. The figure shows that the mere reference to "evidence" decreased considerably, from an initial high of 10% to around 3%. It also shows that the use of a single piece of evidence increased dramatically over the course of the year, from around 13% to nearly 40% of responses. Responses that included two pieces of evidence were roughly stable over the course of the year at around 10%, and those that included three or more pieces of evidence increased from an initial 0% to nearly 9% by the end of the year.

Figure 25 shows change in participants' degree of elaboration of their justifications over the course of the year. Responses that provided no or minimal elaboration, and so failed to show how the reasons cited served to support the relevant model, were roughly stable over the course of the year. However, responses that provided high levels of elaboration, and so provided detailed explanation of the links between the reasons and evidence cited and the relevant model, started the year at 0% and reached over 7% by the end of the year.

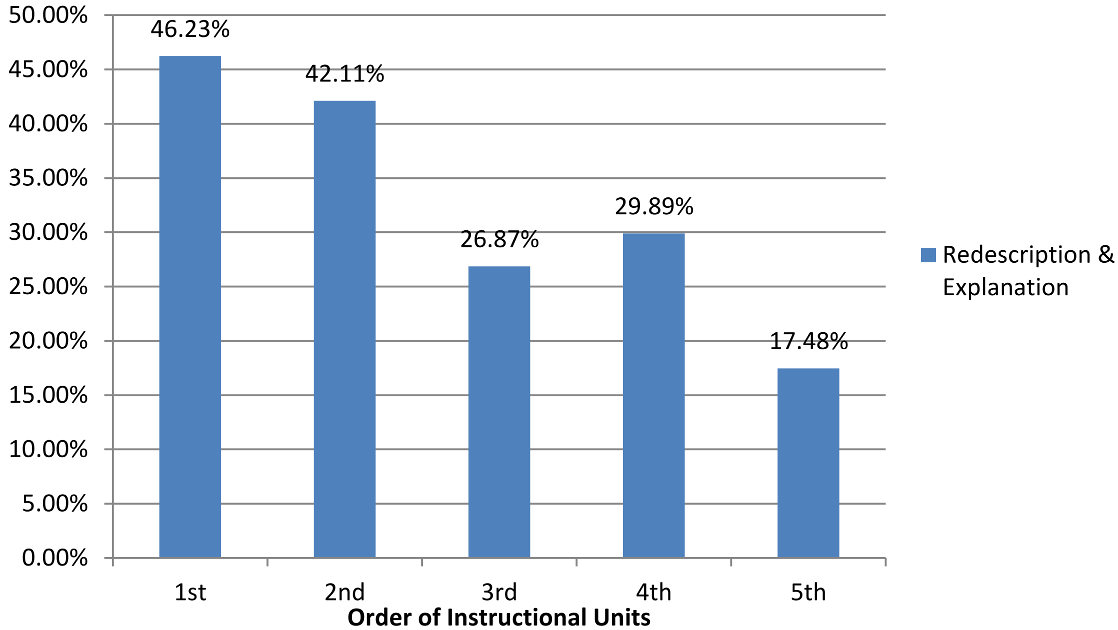


Figure 22. Change in use of non-justificatory “Redescription” and “Explanatory” strategies. This figure shows the changing percentage of responses that merely re-described or re-explained the relevant model.

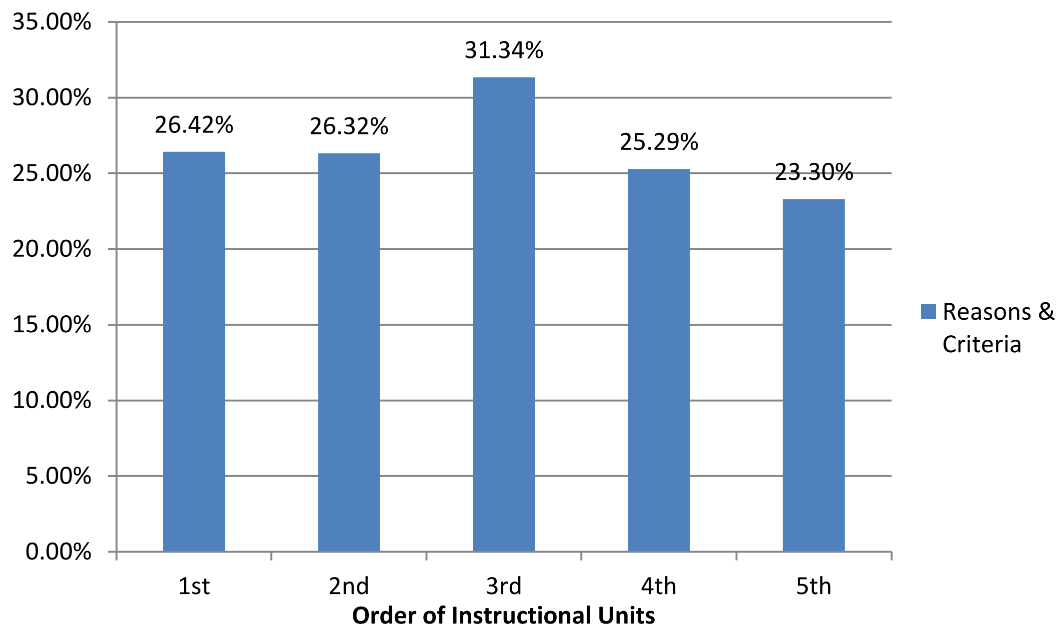


Figure 23. Change in use of “Reasons and Criteria” strategies. This figure shows the changing percentage of responses that mounted justifications using reasons and/or criteria.

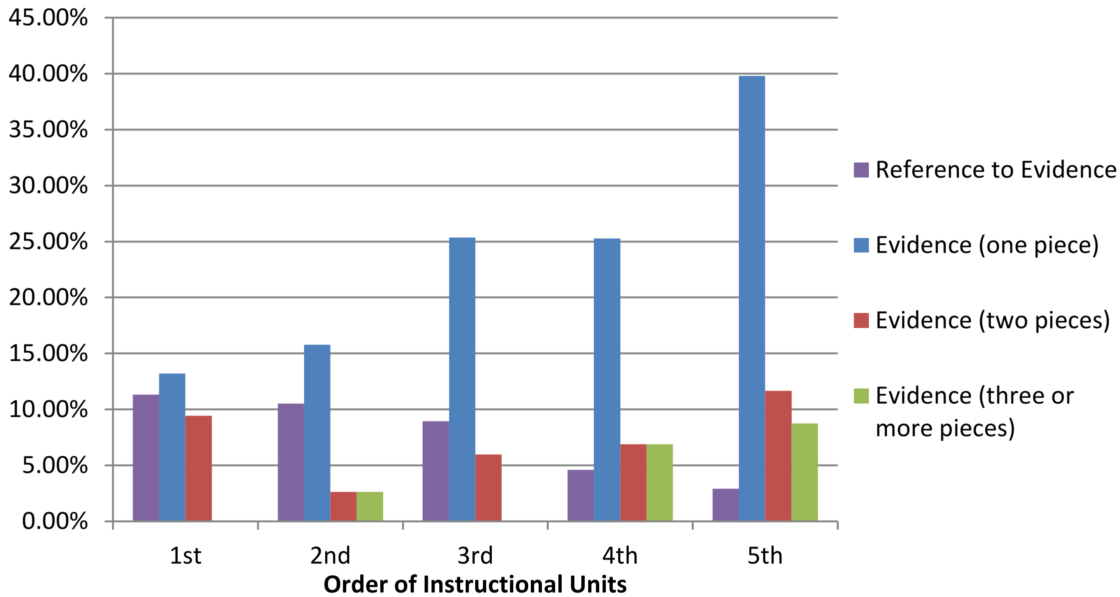


Figure 24. Change in use of evidence across all responses. This figure shows changes in participants use of evidence over the course of the year, capturing responses that merely referred to evidence (without specifying it) , as well as the number of pieces of evidence mentioned.

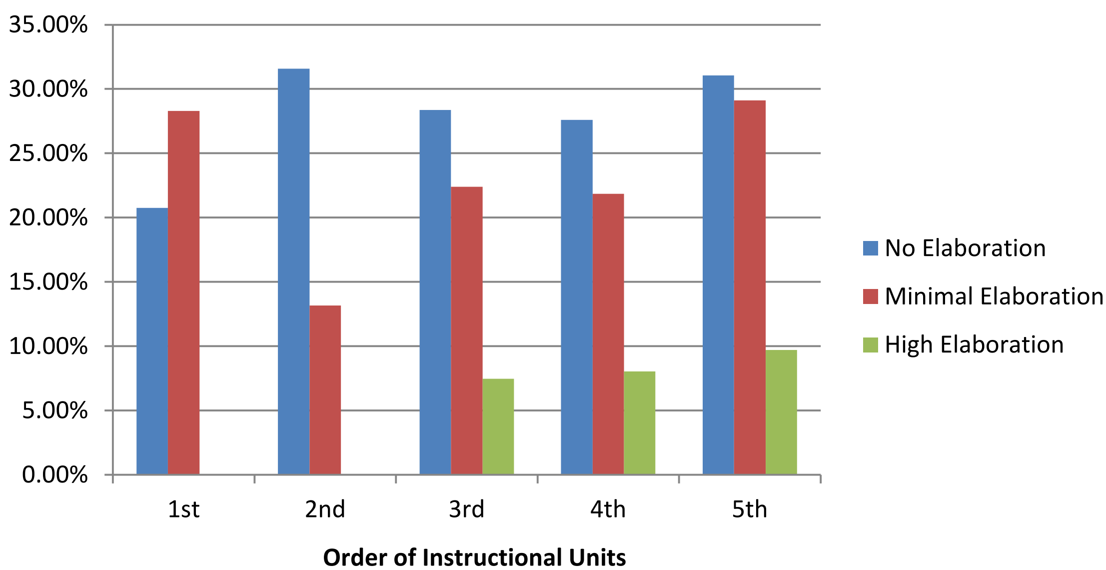


Figure 25. Change in level of justificatory elaboration across responses. This figure shows the changing degree of elaboration exhibited across all responses that mananged to mount a justification (using evidence or criteria).

Finally, Figure 26 shows the change in the relative proportion of the eight kinds of criteria that featured in participants' justifications. *Generic and meta-criteria* featured only in the early parts of the year and at under 5% of criteria-based responses. The same pattern featured for *Collaborative and effortful* criteria, which diminished from around 2% to 0% by year-end. In contrast, *Communicative* criteria featured throughout the year, rising to a high of 15% by the fourth unit and reaching a low of 4.8% by the fifth. *Aesthetic* criteria were generally under 5% of responses, although they declined to 0% of responses by the end of the year. *Evidential* criteria featured in around 10% of all criteria-based responses, and were roughly stable over time, except for the second unit, where they declined to around 5%. *Veridical* criteria varied between 6% and 3%, whereas *Explanatory* criteria decreased from an initial high of 16% to a year-end low of 8%. Finally, criteria of *Logic and Reason* were also roughly stable over the year at around 2%.

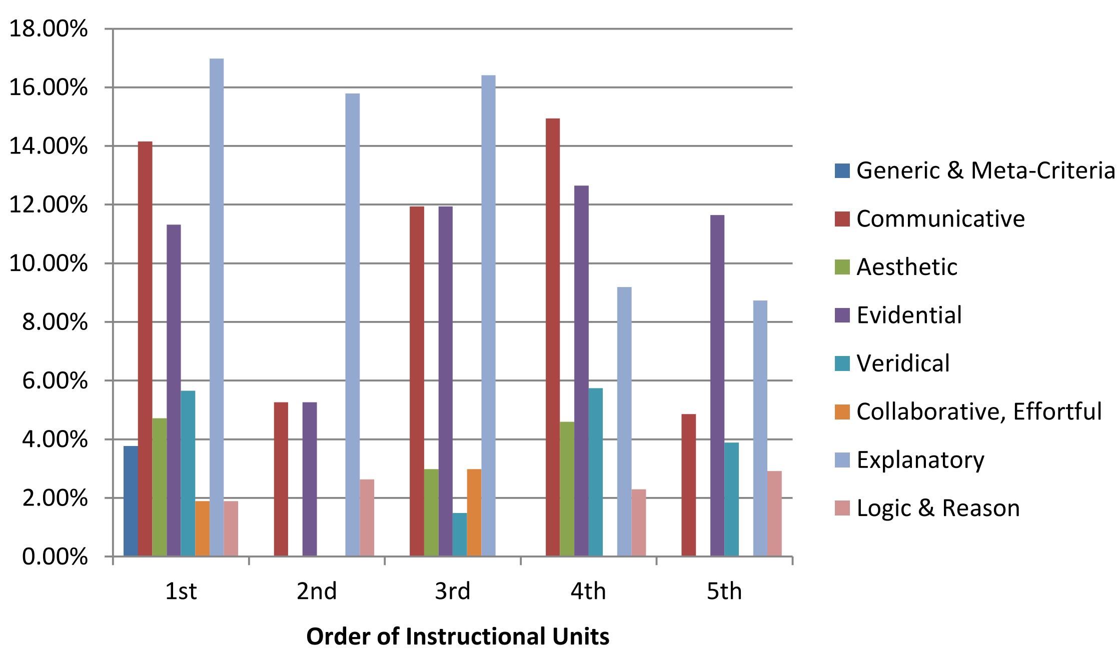


Figure 26. Change in use of modeling criteria across all criteria-based response. This figure shows the changing proportion of criteria that featured in responses that included criteria.

Discussion

Justificatory Strategies and Use of Evidence

Considered overall, a third of all responses expressed reasons and criteria in their justifications, just under half identified specific evidence, and over 10% explicitly identified evidential criteria in their reasoning. This suggests that there was widespread recognition among participants of the need to express reasons, criteria and evidence in their model justifications.

The overall results also reveal some potential limitations in participants' conceptions of justification. Over 14% of responses either failed to specify what particular evidence they invoked in their justification or stated that the evidence was simply "about" the model. This suggests that some participants may have adopted the use of language involving evidence without really understanding the role that evidence plays in justificatory acts. When developing their models participants had at least three independent pieces of evidence available, and usually substantially more than this. That around 10% of responses mentioned two or more pieces of evidence suggests that only some participants recognized the value of using multiple pieces of data in justifying a model. In addition, that 34% of responses simply described or explained the relevant model rather than justifying it, suggests that participants did not always appreciate the essential reason-giving nature of justification.

However, the analysis of change reveals a considerable reduction in participants' reliance on these kinds of non-justificatory strategies over the course of the year. Similarly, the decreasing generic reference to "evidence," the increasing inclusion of evidence as well as of multiple distinct pieces of evidence, all demonstrates an improving awareness of the evidential requirements of good model justifications. In addition, the class criteria generally did not focus on fit with multiple, independent pieces of evidence, and teachers did not make this a focus of instruction. Learners may not have realized the value of including multiple pieces of evidence, even though they considered that evidence in making their judgments.

Crucially, the increase in evidence use and the decrease in non-justificatory kinds of responses occur gradually over the year. This suggests that eliciting these kinds of changes requires constant and extended instruction, which generates incremental improvements in learners' justificatory expertise. Teachers might thus be encouraged to focus on the incremental advancement of learners' conceptions of the nature of justification, building steadily on their previous insights. Instruction might also advance learner conceptions by including explicit discussions on the role of evidence in providing justificatory support for models. This would occur in conjunction with encouragements to include specific evidence. Teachers could also engage learners in discussion about clear examples that distinguish descriptive or explanatory from properly justificatory responses. Familiarizing students with the persuasive power of multiple pieces of independent supporting evidence would also be valuable, and targeted professional development could help to appraise teachers of its importance.

Elaboration of Justification

Although considering evidence and mentioning criteria are vital parts of model evaluation, it is important to determine the ways in which students integrate evidence and criteria into their justifications – particularly because the invocation of evidence might occur in a rote-like or unelaborated manner. As highlighted by Bell and Linn (2000) and McNeill, et al. (2003), good justifications do not merely mention relevant reasons and evidence. Instead, they show in detail how the factors cited actually serve as reasons, by linking them to the claims, explanations, or models being justified. The elaboration codes thus assessed the degree to participants managed to integrate evidence and criteria into their justifications in a meaningful way.

The overall findings show that a third of responses were wholly unelaborated, another third developed minimally elaborated justifications, and 4% were highly elaborated. The analysis of change showed that only the high elaboration category increased over the year, with the other categories remaining roughly stable. That a third of responses involved elaborated reasoning is encouraging, indicating that many participants were able to make meaningful connections between the reasons they cited and the models they justified. The third of responses that failed to develop these links might appear to suggest that there is room for improved instruction focusing specifically on elaborated reasoning.

However, another interpretation of this finding follows from the Berland and Forte (2010) account of the role of the intended target audience in classroom discourse. They have argued that a key component of the design of learning environments is how learners interpret the intended audience of their discourse. One implication of this is that generating more authentic learner argumentation involves finding ways to generate a sense of audience other than the teacher. A fundamental Gricean maxim of conversational implicature, which encapsulates the assumptions that govern the pragmatics of conversation, is the maxim of quantity, in which one aims to provide as much information as needed, and no more (Grice, 1975). In developing their justifications, learners may have tacitly relied on knowledge that they considered as shared with their intended audience – in this case their teacher and peers. Learners may thus have neglected to explain the relation between models and evidence because, given the discursive context of the classroom, this information would have been seen as obvious.

Improving the elaboration of learners' reasoning might therefore require designing materials that make elaborated responses seem necessary in the social context of the classroom. For a start, teachers might begin by focusing learners on their intended audience, and encouraging them to target audiences that lack the relevant background knowledge they attribute to their teachers and peers. Setting and maintaining classroom standards for elaborated verbal discourse might also help learners to practice their powers of elaboration, as would class discussions that explicitly compare highly and poorly elaborated responses. Challenging learners to explain the poorly elaborated justifications of others to their peers could also encourage them to recognize the problems with insufficiently detailed justifications.

Epistemic Criteria for Model Quality

The overall findings show that 30% of responses included criteria for good models in their justifications. In particular, participants most frequently focused on explanatory, communicative and evidential criteria. This focus is plausibly due the efforts of teachers. They frequently encouraged students to avoid focusing on the aesthetic features of their models (e.g. the inclusion of beautiful flowers in models of photosynthesis) and to instead aim for good explanations supported by strong evidence. That only ten percent of responses included evidential criteria belies the far higher rate at which participants mentioned specific evidence. That is, while relatively few responses mentioned evidential criteria, very many recognized the importance of evidence as a powerful source for model justification. Similar considerations are likely to apply to many of the other categories of criteria, with participants recognizing them tacitly without making them explicit in their written justifications.

Although the analysis of change reveals no obvious shifts in participants' preferred criteria over the course of the year, it does show that they cite the full range of modeling criteria from the beginning of instruction. This suggests that middle-school learners generally do not find these kinds of criteria-based judgments involving model quality particularly unfamiliar. They thus do not take very long to adopt and use multiple criteria kinds. The analysis also shows that a wide mix of criteria features in learner justifications throughout the school year, rather than being strongly associated with particular instructional episodes or units.

These results suggest that teachers can productively include multiple criteria from the beginning of instruction, rather than introducing discrete criteria kinds incrementally. However, producing the kinds of shifts observed by Schwarz et al. (2009), which involve a shift in preference from the illustrative to the explanatory functions of models, is likely to require targeted instruction that demonstrates the use of models for facilitating inquiry, rather than merely packaging information.

Conclusions

This study reveals some important strengths in middle-school learners' understanding of the nature and practice of model justification. In particular, it shows their improving conception of justification as involving the giving of reasons in the form of evidence and criteria. In contrast with findings of Jiménez-Aleixandre, Rodríguez, Duschl (2000) and Osborne, Erduran and Simon (2004), and Sadler (2004) that learners' find it difficult to adopt the skills of scientific justification, many participants demonstrated justificatory expertise. In particular, they effectively used evidence to support their preferred models using elaborated reasoning, and showed fluency with various criteria for model quality.

While the results might appear to reveal some important weaknesses in participants' skills of model justification, caution is required in interpreting these findings. First, the Gricean considerations raised above mean that poorly elaborated justifications might not necessarily reveal deficits in participants' capabilities. This is because students' writing typically targets the teacher as an audience, and that these students recognize that the teacher already knows the information involved in the elaborated reasoning. Similarly, participants may have failed to mention evidential criteria because abstract discussions of evidence fit seem superfluous when one cites specific evidence in detail. Participants might thus have omitted important elaborated links and epistemic criteria of which they are fully cognizant.

Second, the study analyzed data drawn from the first year of the PRACCIS instructional intervention. In any large-scale implementation like the PRACCIS project, many procedural and administrative issues require ironing out during early implementation. Given that this was the first year of the project, participating teachers were also just beginning to grapple with the relatively unfamiliar concepts involved as well as with the modeling and criteria-based pedagogical practices of the curriculum. Borko (2004) and Wilson and Berne (1999) have shown that these kinds of professional development interventions involving teacher training are generally difficult and take considerable time and practice. The very steep learning curve faced by learners in adopting new and unfamiliar practices classroom inquiry only compounds this difficulty. Participating teachers were thus far more likely to advance learners' justificatory skills once they have greater familiarity and comfort with the novel concepts and practices of the project. In addition, argumentation instruction was not an explicit focus of instruction in the first year of the PRACCIS implementation, whereas in later years this was a significant focus. For these reasons, the later years of the PRACCIS project involved a more fluid and expert implementation by teachers, promising a greater improvement in learners' justificatory capabilities.

In sum, criteria-centric and model-based approaches to inquiry learning, involving collaborative argumentation using rich bodies of evidence and models represents a promising approach to promoting effective science learning. In particular, they provide a learning environment that encourages the adoption of valuable skills of scientific justification, promote an improved understanding of the evidential and reason-giving nature of justifications, and allow learners to adopt and use sophisticated epistemic criteria of model quality.

Tables

Table 11. Epistemic growth study participants by school, teacher, participant gender and unit.

School	School 1				School 2			
Teacher	Teacher A (male)		Teacher B (female)		Teacher C (female)		Teacher D (male)	
Class	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
Participant Gender	2 male 1 female	2 female 1 male	2 male 1 female	2 female 1 male	2 male 1 female	2 female 1 male	2 male 1 female	2 female 1 male
Order of Instructional Units	1. Cell membranes. 2. Mitosis. 3. Photosynthesis. 4. Cellular respiration. 5. Cardiovascular systems.		1. Cellular membranes. 2. Photosynthesis. 3. Cellular respiration. 4. Mitosis. 5. Genetics.		1. Cellular membranes. 2. Mitosis. 3. Photosynthesis. 4. Cellular respiration. 5. Cardiovascular systems.		1. Cellular membranes. 2. Mitosis. 3. Photosynthesis. 4. Cellular respiration. 5. Cardiovascular systems.	

Table 12. Description of Units

Unit	Description of Unit	Justifications of:
Cellular membranes	Unit explores mechanisms of cell membrane transport, including passive and channel-mediated diffusion and active transport. Inquiry motivated by the problem of lead poisoning; students investigate how lead penetrates human cells.	…student model of iodine diffusion; …student initial model of cell membrane transport; …student model of facilitated diffusion; …group final model of cell membrane transport; …group model comparison; …student model of egg experiment.
Photosynthesis	Unit explores mechanisms and conditions of plant photosynthesis. Inquiry motivated by question of the source and conditions of plant growth, including the role of light, soil, Oxygen, Carbon Dioxide and chloroplasts for purposes of designing a space habitat.	…comparison of competing given models of photosynthesis; …photosynthesis model-evidence diagram judgments; …group final model of photosynthesis.
Cellular respiration	Unit explores mechanisms of cellular respiration. Inquiry motivated by the question of how human bodies use oxygen and food to get energy, with a focus on the processes within cells that facilitate these processes.	…student initial model of cellular respiration; …best student initial model of cellular reparation; …comparison of competing given models of cellular respiration; …cellular respiration model-evidence diagram judgments; …student final model of cellular respiration.
Mitosis	Unit explores mechanisms of cellular growth and replication, with a focus on the processes of mitosis. Inquiry motivated by the question of how living things grow, and how new cells form and change over time.	…student initial model of how things grow; …comparison of competing given models of growth; …comparison of given models of onion cell growth; … student final model of cell growth via mitosis.
Genetics	Unit explores mechanisms of genetic inheritance. Unit focuses on the role of proteins as the link between genes and phenotypic traits. Inquiry motivated by investigation of genetic diseases, including sickle cell anemia and hemochromatosis.	…given model of mechanism of transmission of genetic traits; … student initial model of sickle cell anemia; … comparison of given models of hemochromatosis; …genetics model-evidence diagram judgments;
Cardiovascular systems	Unit explores the cardiovascular system, including the role of the lungs and heart in maintaining living systems. Inquiry motivated by the question of which of various sport activities contributes to cardiovascular fitness most, as well as which matter more for fitness: lung volume or heart rate.	… student initial model of cardiovascular systems and fitness; …cardiovascular system model-evidence diagram judgments; …student final model of cardiovascular systems and fitness; …student revision of final model of cardiovascular systems.

Table 13. Justificatory strategy codes

Code	Explanation	Examples	Prop.
Reasons and criteria	Provides reasons and criteria in support of the model.	"This model is the best one because it really shows you how the Lead gets into the cells by the protien channels."	30.6%
Evidence	Includes any mention of specific described evidence.		47.2%
	Includes mention of one piece of described evidence.	"I think that model 2 is the best because we learned that there isn't much size different between the young and old cell and model 1 says it grows a lot."	36.2%
	Includes mention of two pieces of described evidence.	"The oranges grew more in the CO^2. They got bigger in the glass container. So their experiment did work."	6.1%
	Includes mention of three or more pieces of described evidence.	"My model is good because I used scientific evidence to support it … evidence that the mitochondria make the energy for the cell from our cellular respiration notes. This is important for the model because it is the part of the cell that makes the carbs [sic] into energy. It is where a part of cellular respiration happens. From Figures 4 and 5 … I used the evidence that the more mitochondria and energy makes the person can run faster … It supports my model because it proves that the person with more mitochondria would be ahead. Also … I used the evidence that the cell "burns" glucose with oxygen and produces CO2, energy and water. My arrows from the cell show that they need these to perform cellular respiration. This evidence supports it because it shows that is correct with what the cell needs and produces."	4.9%
Reference to evidence	Refers to evidence, without specifying the evidence.	"The evidence that supports this model is the studies that I've read the stuff shown on the board."	7.1%
Model comparison	Mention and/or critique of alternate models.	"I think it is the best because it has details a neatness, and model #2 and 3 don't."	2.7%
Withhold judgment	Decision withheld (or other epistemic stance expressed).	"I'm not sure. I don't think there is enough evidence."; "Not enough info."	1.0%
Re-description	Restatement, reiteration or re-description of the model.	"Model A says mosquitoes bite small animals like rats, mice, and rabbits. If the mosquitoes then bite a human the red fever virus will transfer to the human."	12.2%
Explanation	Model explanation - no reasons/evidence given.	"I think my model is better because the plastic membrane is probably like the a real membrane and has pours. The iodine is small so it can squeeze through the pours."	21.8%
Non-justificatory	No attempt to provide reasons, explain or describe the model.	"Because that is what I think."	2.2%
About	Claims that reason/evidence and	"Model B talks about bacteria. the penicillin kills bacteria. the evidence and the	7.3%

	model are 'about' the same thing.	model both talk about bacteria somewhat."	
Not interpretable	Response not sufficiently coherent or well-formulated.	"It contradicts Model A because it is about good within supports Model B and not Model A."	0.7%
Blank	Assessment item has not been attempted.	N/A	5.1%

Table 14. Elaboration of justification codes

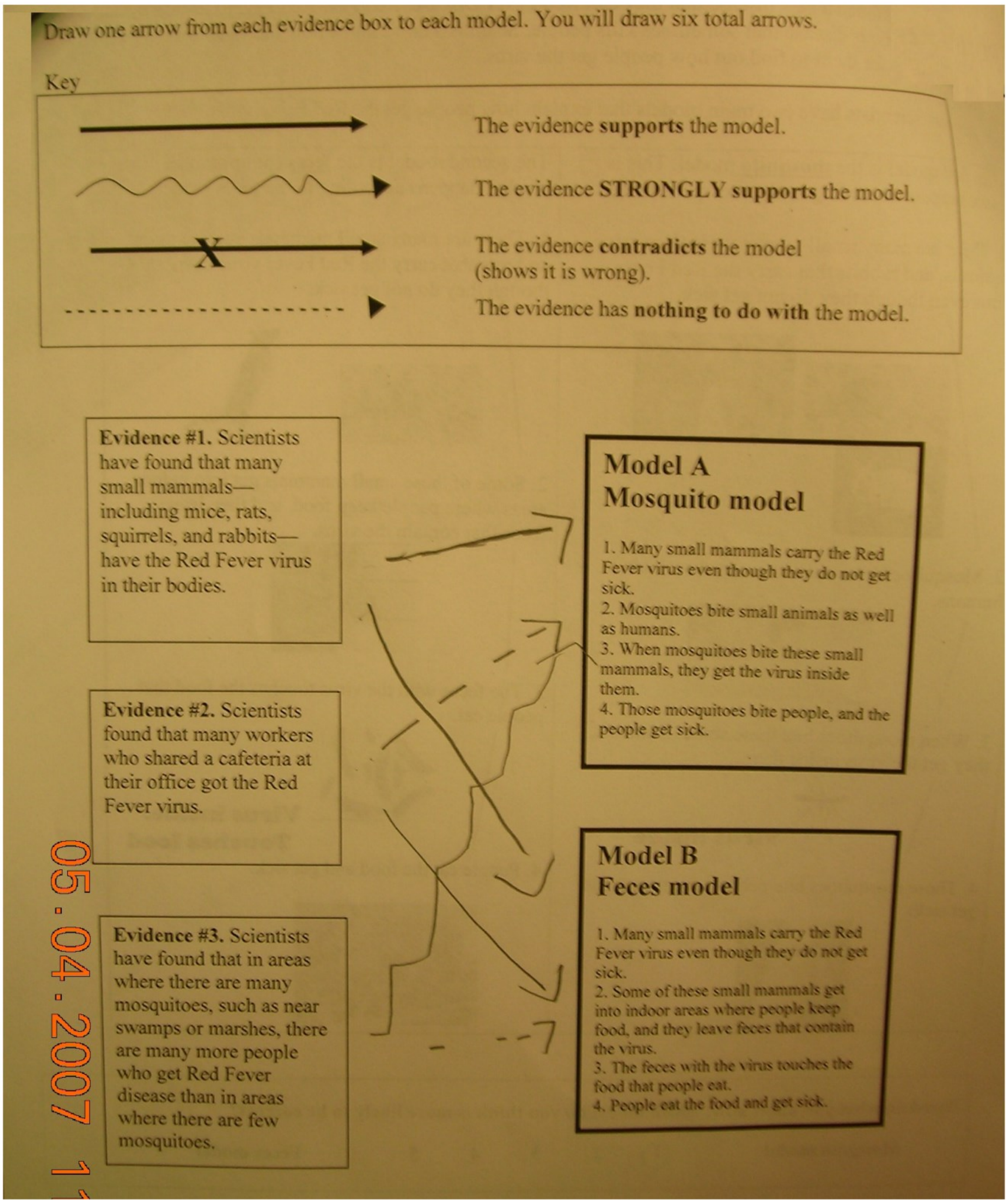
Code	Description	Examples	Prop.
Not applicable	Response provides no reasons, criteria or evidence in support of a model	"Nothing to do with. Does not talk about anything." "It says that they got red fever."	41.6%
No elaboration	Response mentions evidence or criteria in justifying the model, yet provides no substantive link between the model and the reasons or evidence that indicates how or why there is a relation of support	"The evidence states that the plant gets energy from light from CO_2" "Our egg was swollen before I popped it and then the egg looks soggy because it's in the syrup."	29.3%
Minimal elaboration	Response provides a minimal link explaining the relationship between the reason and/or evidence provided with the model. This link shows how the reasons and evidence serve to support the model	"Model 1 says that holes became painful. That evidence says that too. So it strongly supports model 1." "Evidence 2 strongly supports model B because: All the jobs have alot of bacteria intake." "It says that when penicillin is taken to kill bacteria, a persons stomach ulcers disappear much more quickly. This means that it is because the bacteria is no longer living to create the holes." "The model states that the mosquitoes bite the mammals infected with the disease, and the evidence confirms that mammals actually have the disease."	34.2%
High elaboration	Response provides a substantive link explaining the relationship between the reason and/or evidence provided with the relevant model, which is clear and highly detailed	"Ulcers are commonly found in dangerous jobs, like firefighter, or coal miners. Those types of jobs can be hard on the body causing stress that leas to excessive body acid which causes ulcers." "Model B says that people get the red fever virus from infected food, and if people share the same cafeteria they eat the same food, which means if the food was contaminated they would all eat it, and get sick. I do not think it strongly supports it because if they all went to the same cafeteria they would all live in the same area and they could get it from bugs (model A)."	4.4%

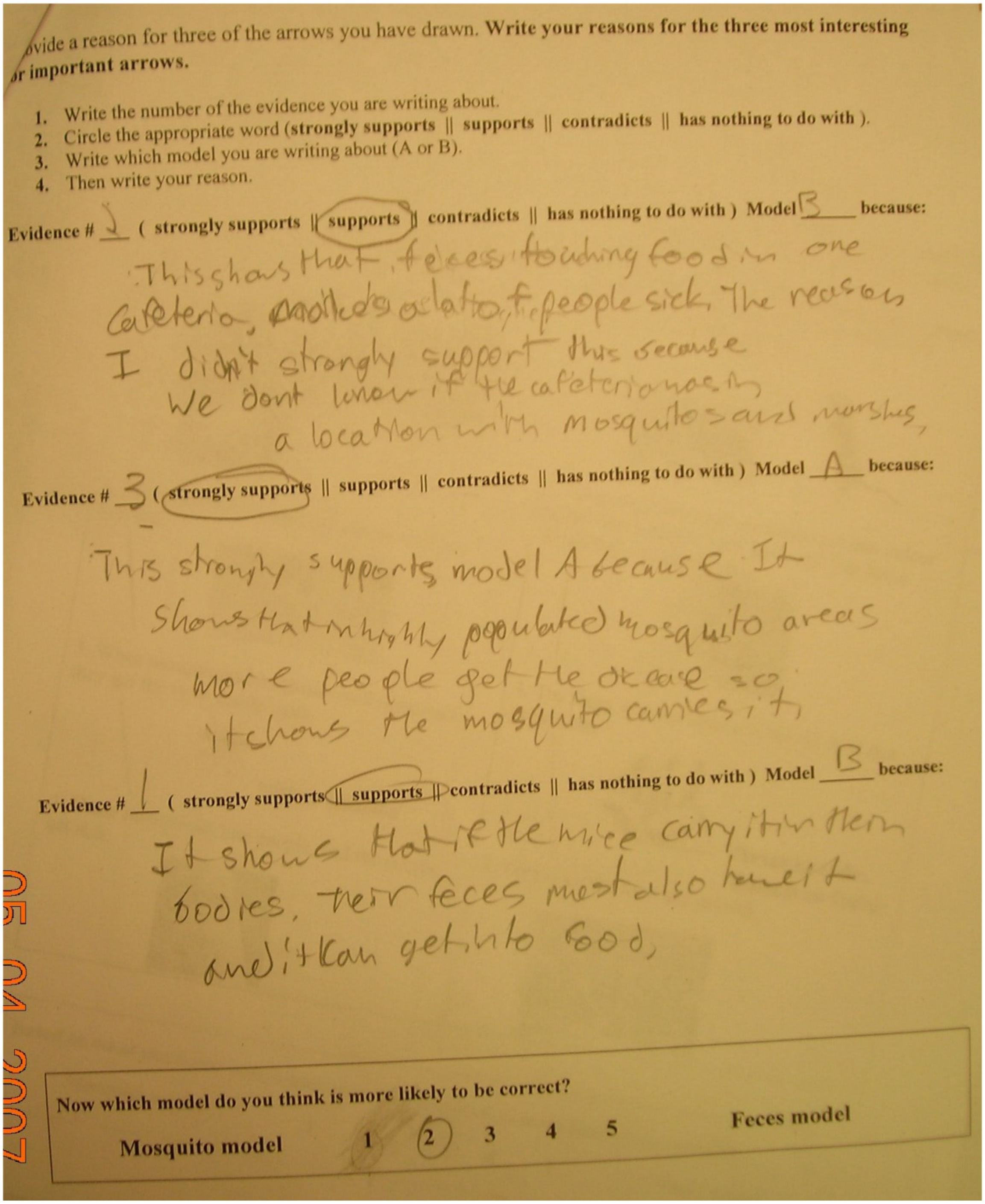
Table 15. Categories of modeling criteria codes

Criteria type	Description of model	Examples	Prop.
Generic & meta-level	Model is good, better or as fulfilling unspecified "criteria."	"I think they should add this information into their model because it may change their results and make them better."	1.5%
Communicative	Model has characteristics that aid in the communication of ideas to the reader.	"My model is the best because it is clear and easy to understand, incorporates all the evidence we found out about photosynthesis."; "… Everything is obvious."	13.7%
Aesthetic	Model has specific aesthetic features.	"I think it is the best because it has details a neatness, and model 2 and 3 don't"; "…It also has color and it catches your eyes not boring like one color."	1.7%
Evidential	Model has a degree of support by evidence or facts.	"Our model is the best because we have many sources to justify it."; "I think it is the best because it is the one model I know the most about…"	11.5%
Veridical	Model is good/better in virtue of it being true or accurate.	"We think "model C" is best because it has the most evidence to support it. It is the only model that correctly refers to protein. It also has a good sequence of events."	3.9%
Explanatory	Model is good/better in virtue of its explanatory, representative and descriptive functions: it reveals how something happens.	"The bag has little holes and the cells are so small they can get through. I think that my models looks better because it is explaining that the cell membrane has holes but so does the bag. So the cells can just squeeze through the hole in the bag, and my drawing explains that. "	16.4%
Logic & reason	Model is logical or sensible/reasonable.	"It seems to make more sense. The growth model does not work because all cells are microscopic."	2.2%
Collaborative & Effortful	Model is the product of the work of several people.	"I think my model is a good one because my whole group helped make it…"	1.0%
None	No criteria are mentioned.	"Glucose + Oxygen are important for muscle contraction and cellular respiration. During cell respiration process, CO_2 is released from the cell and we exhale it."	59.9%

Appendices

Appendix A. Example of a model-evidence diagram and associated justifications





ovide a reason for three of the arrows you have drawn. **Write your reasons for the three most interesting or important arrows.**

1. Write the number of the evidence you are writing about.
2. Circle the appropriate word (**strongly supports || supports || contradicts || has nothing to do with**).
3. Write which model you are writing about (A or B).
4. Then write your reason.

Evidence # 2 **(strongly supports || supports || contradicts || has nothing to do with) Model** B **because:**

This shows that feces touching food in one cafeteria, made a lot of people sick. The reason I didn't strongly support this because we dont know if the cafeteria was in a location with mosquitos and marshes,

Evidence # 3 **(strongly supports || supports || contradicts || has nothing to do with) Model** A **because:**

This strongly supports model A because It shows that in highly populated mosquito areas more people get the disease so it shows the mosquito carries it,

Evidence # 1 **(strongly supports || supports || contradicts || has nothing to do with) Model** B **because:**

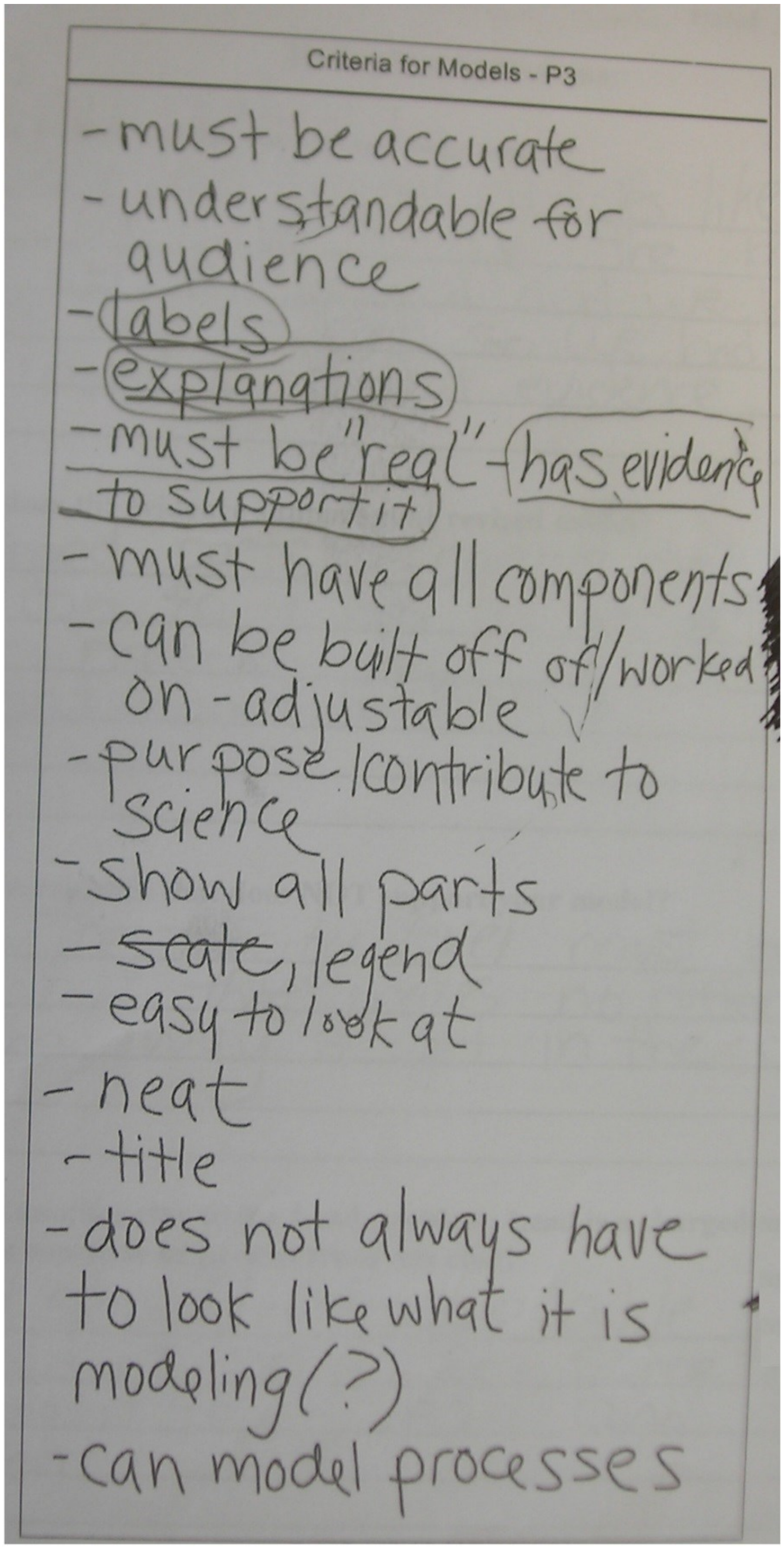
It shows that if the mice carry it in them bodies, their feces must also have it and it can get into food,

Now which model do you think is more likely to be correct?

Mosquito model 1 (2) 3 4 5 **Feces model**

05 04 2007

Appendix B. Examples of a student's epistemic criteria for model quality



Criteria for Models - P3

- must be accurate
- understandable for audience
- labels
- explanations
- must be "real" - has evidence to support it
- must have all components
- can be built off of/worked on - adjustable
- purpose/contribute to science
- Show all parts
- ~~scale~~, legend
- easy to look at
- neat
- title
- does not always have to look like what it is modeling (?)
- can model processes

References

Bell, P., & Linn, M. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education*. 22 (8), 797-817.

Berland, L. K., & Forte, A. (2010). When students speak, who listens? Constructing audience in classroom argumentation. In *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1* (pp. 428-435). International Society of the Learning Sciences.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765-793.

Berland, L. K., & McNeill, K. L. (2012). For whom is argument and explanation a necessary distinction? A response to Osborne and Patterson. *Science Education*, 96(5), 808-813.

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

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3-15.

Bransford, J. D., A. L. Brown, et al., Eds. (2000). *How People Learn: Brain, Mind, Experience and School*. Washington D.C., National Academy Press.

Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education*, 92(3), 473-498.

Buckland, L. A., & Chinn, C. A. (2010). Model-evidence link diagrams: A scaffold for model-based reasoning. In *Proceedings of the 9th International Conference of the Learning Sciences*, Volume 2, (449-450). International Society of the Learning Sciences.

Buckland, L.A., & Chinn, C.A. (2015). Epistemology. *Encyclopedia of Science Education*. Gunstone, R. (Ed.) Springer Press.

Buckley, B. C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education*, 22(9), 895-935.

Chinn, C.A. & Buckland, L.A. (2011). Model-based instruction: Fostering change in evolutionary conceptions and in epistemic practices. In Rosengren, K. S., Evan, E. M., Brem, S., & Sinatra, G. (Eds.), *Evolution Challenges: Integrating research and practice in teaching and learning about evolution.* Cambridge: Oxford University Press.

Chinn, C.A., Buckland L.A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: arguments from philosophy and psychology, *Educational Psychologist*, 46, 141–167.

Chinn, C.A., Duschl, R.A., Duncan, R.G. & Buckland, L.A. (2008). A microgenetic classroom study of learning to reason scientifically through modeling and argumentation. In *ICLS 2008: Proceedings of the International Society of the Learning Sciences*. Lulu.

Chinn, C.A., Pluta, W.J., Buckland, L.A., Rogat, T.K., Difranco, J., & Witham, S. (2010). Promoting reasoning: A microgenetic study of middle-school students learning through model-based inquiry. Symposium presented at the annual meeting of the American Educational Research Association, Denver, CO.

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In D. Rapp and J. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences.* (pp. 425-453). Cambridge, MA: MIT Press.

Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, 22, 1041–1053.

Coll, R. K., France, B., & Taylor, I. (2005). The role of models/and analogies in science education: implications from research. *International Journal of Science Education*, 27(2), 183-198.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, 32, 268–291.

Duschl, R., & Osborne, J. (2002). Supporting and promoting argumentation: Discourse in science education. *Studies in Science Education,* 38, 39-72.
Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press.
Giere, R.N. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.
Giere, R.N. (2004). How models are used to represent reality. *Philosophy of Science*, 71, 742–752.
Gobert, J. (2000). A typology of models for plate tectonics: Inferential power and barriers to understanding. *International Journal of Science Education*, 22(9), 937-977.
Gobert, J. D., & Pallant, A. (2004). Fostering Students' Epistemologies of Models via Authentic Model-Based Tasks. *Journal of Science Education and Technology*, 13, 7-22.
Gobert, J. D., & Buckley, B. C. (2000). Introduction to model-based teaching and learning in science education. *International Journal of Science Education*, 22(9), 891-894.
Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, 21, 725–740.
Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, 43(3), 142-160.
Grice, H. (1975). Logic and Conversation. *Syntax and Semantics*, *3*, 41-58.
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, 799–822.
Gutwill, J.P., Frederiksen, J.R., & White, B.Y. (1999). Making their own connections: Students' understanding of multiple models in basic electricity. *Cognition and Instruction*, 17, 249–282.
Harrison, A.G., & Treagust, D.F. (2000). Learning about atoms, molecules and chemical bonds: A case study of multiple-model use in grade 11 chemistry*. Science Education*, 84, 352–381.
Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. NJ: Erlbaum: Mahwah.
Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching*, 38, 663–687.
Jiménez-Aleixandre, M. P., Rodríguez, A. B., & Duschl, R. A. (2000). “Doing the lesson” or “doing science”: argument in high school genetics. *Science Education*, 84, 757-792.
Kelly, G. J. (2008). Inquiry, activity, and epistemic practice. In R. A. Duschl & R. E. Grandy (Eds.), *Teaching scientific inquiry: Recommendations for research and implementation*. (99–117). Rotterdam, The Netherlands: Sense Publishers.
Kitcher, P. (1993). *The advancement of science.* Oxford: Oxford University Press.
Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, 92, 1-32.
Kuhn, T.S. (1977). *The essential tension*. Chicago: University of Chicago Press.
Kuhn, D. (1991). *The skills of argument.* Cambridge, UK: Cambridge University Press.
Lederman, N. (1992) Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29 (4), 331-59.
Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In R. Keith Sawyer (Ed.), *Cambridge handbook of the learning sciences*, 371–387, Cambridge: Cambridge University Press.
McNeill, K. (2011). Elementary students' views of explanation, argumentation, and evidence and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, 48(7), 793–823.

McNeill, K. L., & Krajcik, J. (2012). S*upporting grade 5-8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing.* New York, NY: Pearson Allyn & Bacon.

McNeill, K. L., Lizotte, D. J, Harris, C. J., Scott, L. A., Krajcik, J., & Marx, R. W. (2003). Using backward design to create standards-based middle-school inquiry-oriented chemistry curriculum and assessment materials. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences*, 15(2), 153–191.

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology*, 33, 177–208.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Edwards, M. N. (2007). Epistemological threads in the fabric of conceptual change research. In S. Vosniadou, A. Baltas, & X. Vamvakoussi (Eds.), *Reframing the conceptual change approach in learning and instruction*. 105–122. Oxford, UK: Elsevier .

Nagel, E. (1961). *The structure of science: Problems in the logic of science education.* New York, NY: Harcourt, Brace, & World, Inc.

Nersessian, N. (2002). The cognitive basis of model-based reasoning in science. In P. Carruthers, S. Stich, & M. Siegal (Eds.). *The cognitive basis of science.* 17–34. Cambridge: Cambridge University Press.

National Science Teacher Association (2013). *Next generation science standards*. Retrieved from www.nextgenscience.org

Osborne J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science,* 328, 463–466.

Osborne, J., (2014). Scientific practices and inquiry in the science classroom. In Lederman, N. G., & Abell, S. K. (Eds.). *Handbook of Research in Science Education* (Vol. 2). Routledge.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994-1020.

Osborne, J., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, 95(4), 627–638.

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-Based Inquiry and School Science: Creating Connections. *School Science and Mathematics*, *109*(7), 394-402.

Penner, D.E., Giles, N.D., Lehrer, R., & Schauble, L. (1997). Building functional models: Designing an elbow. *Journal of Research in Science Teaching*, 34. 125–143.

Pluta, W. J., Buckland, L. A., Chinn, C. A., Duschl, R. A., & Duncan, R. G. (2008). Learning to evaluate scientific models. In *ICLS 2008: Proceedings of International Society of the Learning Sciences*. Lulu.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, 48(5), 486-511.

Raghavan, K., & Glaser, R. (1995). Model-based analysis and reasoning in science: The MARS curriculum. *Science Education*, 79, 37–61.

Rosebery, A., Ogonowski, M., DiSchino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *Journal of the Learning Sciences*, 19, 322–357.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*. 41(5). 513-536.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45, 1–32.

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

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

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., & Cam, A. (2011). Elementary children's judgments of the epistemic status of sources of justification. *Science Education*, 95(3), 383-408.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632-654.

Schwarz, C.V., & White, B.Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, 23, 165–205.

Schweingruber, H., Keller, T., & Quinn, H. (Eds.). (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.

Smith, C.L., Maclin, D., Houghton, C. & Hennessey, M.G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction*, 18, 349–422.

Songer, N. B., & Gotwals, A. W. (2012). Guiding explanation construction by children at the entry points of learning progressions. *Journal for Research in Science Teaching*, 49, 141–165.

Toth, E.E., Suthers, D.D., & Lesgold, A.M. (2002). ''Mapping to know'': The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education*, 86, 264–286.

Toulmin, S. (1958, 2006). *The uses of argument* (updated edition). New York: Cambridge University Press.

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of scientific models in learning science. *International Journal of Science Education,* 24, 357–368.

White, B.Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10, 1–100.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16, 3-118.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of research in Education*, 173-209.

Windschitl, M., Thompson, J., & Braaten, M. (2008). How novice science teachers appropriate epistemic discourses around model-based inquiry for use in classrooms. *Cognition and Instruction*, 26, 310–378.

CONCLUSION

The three studies presented above explore some of the disparate ways in which the expanded framework for epistemic cognition developed in Chinn et al. (2011) provides a rich resource for new ways of thinking about cognition, learning, and inquiry. Each study uses either the framework itself (Chapter 1), a component of the framework (Chapter 2), or the recommendations developed in the article for advancing research into epistemic cognition (Chapter 3).

The first study surveyed the conceptions addressed over the history of NOS research, using the expanded framework to make sense of the scope of conceptions that NOS instruments have addressed. The resulting analytic framework traces the scope of the topics that feature in a diverse range of 81 NOS instruments used in six decades of research. The analysis of NOS research reveals considerable change in the kinds of conceptions targeted for investigation in the field, and it does so at a considerably higher level of detail while including far more instruments than has extant research.

In particular, the analysis reveals a surge in the number of NOS studies from 1996, producing an increasingly rich array of science conceptions investigated, and dramatic growth in the use of open-ended survey items. Change has also occurred in the focus on non-doxastic, attitudinal measures, which target scientific attitudes and beliefs about science learning, rather than propositional beliefs about the nature of science. While by the late 1970s, NOS research generally excluded attitudinal constructs, the large-scale surveys of the 1980s reincorporated measures targeting scientific attitudes. In addition, over the long term, the field has seen a decline in the focus on scientific aims, epistemic virtues and vices, and on non-epistemic descriptors of science, scientific aims and scientists.

Many of the observed changes in NOS instrumentation are likely to be the result of changes in the views of science held by philosophers, society and NOS researchers themselves. Several large-scale reviews of the field of NOS inquiry have developed hypotheses about these kinds of influences. For example, the current study provides mixed support for claims of a Kuhnian transition proposed by Abd-El-Khalick and Lederman (2000). On the one hand, the finding that early instruments featured a comparatively outsized focus on social features of science contradicts the claim. On the other hand, the increasing focus on scientific sources, certainty and change, disagreement and rival conceptions, as well as on the theory-laden nature of science supports the idea of a Kuhnian change in NOS research.

The expanded framework for epistemic cognition defended in Chinn et al. (2011) thus provides a useful tool for understanding the scope of research into people's scientific epistemologies. It allows for the mapping of scientific epistemic cognition over a broad array of issues while addressing specific conceptions at a fine grain-size. In providing detailed insight into variation in the conceptions of NOS that feature in the field of research, the framework provides a valuable resource for the design of future NOS instruments. In particular, the analysis reveals that the field of NOS instrumentation as neglects some important components of the expanded framework. For example, NOS instruments typically include few items targeting conceptions involving scientific aims, particularly epistemic aims, and very few targeting the epistemic value of science. It also includes almost no items on the epistemic responsibilities of scientists and non-scientists. These neglected areas correspond to important components of the expanded framework, and each could thus be more intensively investigated in future NOS work.

The second study follows the recommendations expressed in Chinn et al. (2011), investigating a considerably under-researched aspect of epistemic cognition – beliefs about reliable processes of knowledge production. The interview and written data analyzed reveals participating undergraduates' had epistemic beliefs about a wide array of processes and their conditions of reliability. These included perceptual, cognitive, social, testimonial, institutional, and evidential processes. Participants exhibited considerable variation in the range and kinds of processes and conditions they considered relevant to the production of knowledge during their spontaneous, unguided reasoning. Some participants mentioned conditions on processes of knowledge production at twice the rate of others, and some mentioned four times as many kinds of conditions on processes. This shows that some undergraduates have far richer conceptualizations of knowledge production than their peers, a difference to which existing measures of epistemic cognition would be largely insensitive.

During guided episodes of reasoning, participants were typically able to recognize and evaluate a wide array of processes of knowledge production. This suggests that while some might battle to identify important processes spontaneously, directed reflection can activate a complex web of beliefs involving knowledge production that people already possess. This widespread facility with notions of knowledge production has positive implications for process-oriented instruction, suggesting that learners are likely to engage readily in collaborative argumentation about processes, and to use the vocabulary of processes in their evaluations of knowledge attributions. However, in spite of their facility with notions involving knowledge production, the short period spent in directed inquiry did not detectably affect participants' reliance on these notions during a subsequent written analysis of a complex and authentic case. This suggests that more intensive instruction is likely required to get learners to activate the beliefs they hold about processes in their reasoning.

Participants also demonstrated understanding of processes at different levels of sophistication, demonstrating various conceptual strengths and weaknesses. The analysis reveals deficits in participants' understanding of the epistemic basis of knowledge derived from eyewitness testimony. In particular, they neglected to consider the importance of the conditions of reliability on the cognitive and institutional processes this knowledge involves. They also tended to neglect important social and testimonial features of collaborative inquiry, particularly the deliberative procedures followed by inquirers and the flow of information amongst them. Again, existing measures are unlikely to detect these features of undergraduates' reasoning about knowledge production.

These findings thus challenge the theoretical orthodoxy that dominates the field of research into epistemic cognition, which largely neglects these kinds of beliefs. They also indicate that this component of the expanded framework constitutes a valuable target for future research and instruction. As beliefs about processes of knowledge production are vital for evaluating claims and attributions of knowledge, further research could productively incorporate them into models and measures of epistemic cognition. A useful next step would be to explore relationships between the beliefs about processes and orthodox measures. More generally, these findings also recommend the utility of the expanded conception of epistemic cognition explicated in Chinn et al. (2011).

Finally, the third study explores the epistemic criteria implicit in the model-based justifications of 24 seventh-grade students engaged in inquiry learning. It thus advances our understanding of how learning environments featuring epistemic criteria for model-based inquiry can promote participants' skills of justification. Participating students exhibited a complex array of higher-level criteria in their judgments about model quality, and varied considerably in their justificatory expertise.

In particular, participants exhibited widespread recognition of the need to express reasons, criteria for model quality and evidence in their justifications, with half of all justifications including specific evidence. A third of all justifications developed minimally elaborated descriptions of the relationship between the reasons cited and the relevant model, with just under 5% providing highly elaborated justifications. A third of all responses featured identifiable epistemic criteria for model quality, with explanatory, communicative and evidential criteria dominating, in addition to aesthetic criteria, veridical criteria and criteria of logic and reason.

In addition, the analysis of change revealed a general reduction in participants' reliance on non-justificatory strategies, like the mere description or explanation of the relevant model. They also showed that participants' justifications increasing relied on evidence and multiple pieces of evidence. While levels of elaborative sophistication were generally stable, the high elaboration category also increased over the course of the year.

The study thus provides pedagogically valuable insight into the characteristic strengths and weaknesses of science learners' model-based justificatory practices. The slow and steady nature of the observed changes implies that eliciting them requires constant and sustained instruction that generates incremental improvements in skill. Improving skills of elaboration is also likely to require materials and interactions that make highly elaborated justifications seem necessary and natural in the social context of the classroom. That participants were able to cite the full range of modeling criteria from early in the year demonstrates that young learners are familiar with these kinds of evaluative concepts, and readily adopt them into their classroom reasoning.

The third study thus demonstrates another way in which the Chinn et al. (2011) article advances the field of research. It does so by following the recommendation outlined in the article to investigate the role of epistemic criteria in justification. Further research could investigate the kinds of changes engendered in justificatory skill by using epistemic criteria over a longer period of learning, across multiple disciplines, and with expert teachers.

Each of these three inter-related studies reveals the fruitfulness of an expanded conception of epistemic cognition. The first demonstrates the applicability of the expanded framework for the fine-grained analysis of research instrumentation targeting people's scientific epistemologies. The second study reveals the little-recognized importance of beliefs about reliable processes of knowledge production for understanding claims and attributions of knowledge. The third study reveals change in the justificatory practices of middle school learners engaged in sustained criteria-centric and model-based science learning through inquiry. Each of the three studies thus shows how expanded notions of epistemic cognition can advance our understanding of this psychologically and pedagogically vital aspect of cognition.