

Coherence vs. Fragmentation in student epistemologies: A reply to Smith & Wenk

Andrew Elby
University of Maryland

Abstract

In the literature on conceptual change and students' intuitive preconceptions about the natural world, a long-running paradigm debate pits the Coherence perspective against the Fragmentation perspective. The Coherence perspective attributes stable, robust, context-independent concepts or intuitive theories to students, while the Fragmentation perspective describes students' reasoning in terms of the context-dependent activation of finer-grained knowledge elements that are more loosely organized than "theories." Over the past few years, this paradigm debate has entered the literature on students' personal epistemologies, their views about the nature of knowledge and knowing. Epistemology researchers have generally assumed student epistemologies to consist of robust, globally applicable cognitive structures — epistemological theories, belief systems, or stages — that fit into the Coherence perspective. However, cognitive frameworks that fit largely or wholly into the Fragmentation camp recently began to challenge this consensus. Smith & Wenk (2006) published the first large-N study designed to experimentally distinguish the Coherence and Fragmentation perspectives. Smith & Wenk conclude that their data favor Coherence over Fragmentation. With the emergence of this paradigm debate in the epistemology literature, a crucial methodological issue becomes, What counts as evidence for Coherence at the expense of Fragmentation, or vice versa? This article attempts to further the scholarly discussion of this issue. To do so, I use a critique of Smith & Wenk's study as a launching point from which to make more general methodological arguments.

Correspondence concerning this manuscript should be addressed to Andrew Elby, Dept. of Curriculum & Instruction, University of Maryland, College Park, MD 20742-1115

Introduction

In the literature describing people's intuitive conceptions of the natural world, a growing body of research attempts to address the paradigm debate between two descriptions of novices' conceptual reasoning: the *Coherence* perspective, which attributes stable, robust, comparatively context-independent concepts or intuitive theories, and the *Fragmentation* perspective, which models student reasoning in terms of the context-dependent activation of finer-grained knowledge elements that are more loosely organized than "theories" (Anderson, Tolmie, Howe, Mayes, & Mackenzie, 1992; Clark, 2003; diSessa, Gillespie, & Esterly, 2004; Elby, 2000; Ioannides & Vosniadou, 2002; Samarapungavan & Wiers, 1997). By contrast, in the literature addressing epistemologies — students' views about the nature of knowledge and knowing (Hofer & Pintrich, 1997, 2002) — almost no experimental work has pitted the Coherence and Fragmentation perspectives against one another.

A rare exception is Smith & Wenk's (2006) study, which used three different probes of students' views about the nature of scientific knowledge. Predicted patterns of epistemological coherence across the three probes, they argue, would count as evidence that students possess a robust, comparatively context-independent cognitive structure called an *intuitive meta-theory*, a construct that fits into the Coherence perspective. By contrast, epistemological inconsistency across the three probes could favor the Fragmentation perspective. Smith & Wenk claim that their data support Coherence.

In this article, I use Smith & Wenk's paper as a launching point to make methodological arguments about what counts as evidence for Coherence at the expense of Fragmentation in personal epistemology research. These arguments can inform the paradigm debate as experimental tests become more common. To present and illustrate my points in a concrete way, I argue that Smith & Wenk's article does not support Coherence at the expense of Fragmentation, for three reasons. First, the data and analysis they present are insufficient to establish that individual subjects displayed the same (or approximately the same) level of epistemological sophistication across the three probes. Second, even if subjects were epistemologically consistent across the three probes, *priming* and *contextual stability* can account for that coherence, with no need to attribute theory-like cognitive structures to subjects. Third, Smith & Wenk's three probes are sufficiently similar that, to students, they might not constitute distinct contexts; and the Fragmentation perspective can account for *local* coherences — epistemological coherence within a context — just as well as the Coherence perspective can. Therefore, epistemological consistency across the three probes would not favor the Coherence perspective. I end by posing a variant of Smith & Wenk's experiment that, I argue, avoids these methodological problems and constitutes a fair test of Coherence vs. Fragmentation.

Smith & Wenk's experiment and argument

Smith & Wenk (2006) is to my knowledge the first personal epistemology research that attempts an experimental test of Coherence vs. Fragmentation. Crisply summarizing the current paradigm diversity in the field, the authors write

Although there is broad agreement that developing students' epistemological understandings may be important as an educational outcome, there is less consensus about how these understandings should be conceptualized and assessed. One issue is the extent to which epistemological views are "tacit" versus "articulated." Other issues concern what "grain size" it is useful to use in describing these understandings, the ways that context affects them, the extent to which even the novice's thinking (as well as the expert's) is coherent and organized, and the extent to which epistemological understandings are domain general versus domain specific. At present, researchers have conceptualized student epistemological understandings as a domain-general developmental structure (King & Kitchener, 1994), as domain-general beliefs (Schommer, 1990), as domain-specific theories (Hofer,

2004; Hofer & Pintrich, 1997), and as highly situation-specific resources (Hammer & Elby, 2002). We believe that debate about these issues can be productive at this early point in charting out this complex terrain, especially if those with competing views take care to articulate their positions clearly, generate testable predictions from their views, and identify the important phenomena about student thinking and reasoning that any theory of epistemological reasoning needs to account for. Further, it might even be the case that each tradition is trying to understand an important “part of the elephant” and that we may ultimately have more comprehensive theories that integrate diverse views. (p. 748)

To test the Coherence perspective (developmental structures or theories) against the Fragmentation perspective (resources), Smith & Wenk

assess epistemology using three different types of probes.... Consistency among multiple probes is one important way to assess the validity of epistemology of science as a coherent construct. That is, if it is a coherent or intuitive theory, then the various aspects of epistemology are facets of a system of thinking, and one would expect interrelationships among them. Of course, this assumption is open to disconfirmation if the different facets are independently scored as in this study. One could find that there was little within-student consistency across different probes. (p. 752)

Smith & Wenk’s three probes are three parts of a 45-minute clinical interview conducted with each of 35 subjects.

The three parts of Smith & Wenk’s interview

Part 1. This part is a modified version of the Carey, Evans, Honda, Jay, & Unger (1989) and Smith, Maclin, Houghton, & Hennessey (2000) Nature of Science interview consisting of 7 questions, some with multiple parts. They include

- 2a. What sorts of questions do scientists try to answer? Can you give me a specific example of a question that a scientist would ask? 2b. What would a scientist do to answer that question? Would they do experiments? Can you give me an example of an experiment they might do?...
- 4a. What is a scientific hypothesis? Can you give a specific example of a scientific hypothesis? 4b. Does a scientist’s hypothesis influence the experiments (s)he does?...
- 6a. Are theories and hypotheses related? If yes, how? 6b. Does a scientist’s theory influence the hypotheses he or she tests in a particular experiment? If yes, how? Can you give an example of that? If no, why not?

Part 2. Subjects are presented with the following statement:

“Every day, in more and more areas of science the right answer is known. In areas where the right answer is known, I look to experts to tell me what is right. In areas where no right answer is known, I think anyone’s opinion is as good as another’s.”

The subjects then respond to these questions:

- a. Do you agree with this statement? Why or why not?
- b. Do you think science knows the right answers? Why or why not?
- c. Where do you go when you have questions about a scientific issue?
- d. What do you do when you find disagreement among sources?
- e. Do you agree with this person who says that when there are no right answers anybody’s opinion is as good as another’s?
- f. Is there anything that makes one answer better than another?

Part 3. Students are told, “I will be presenting a situation and asking you some questions about it. I am not concerned with how much information you have about the issues, but how you think about them...” Some subjects then read and hear a controversy about water fluoridation, while others read and hear the following:

The Effects of Echinacea. Preparations of Echinacea root are among the most popular herbal supplements in the United States marketplace. Echinacea is used for preventing and treating the common cold, flu, and upper respiratory tract infections. It’s also used to increase general immune system function. Because Echinacea has become so popular, scientists have been conducting tests on the effectiveness of Echinacea. Some of the studies show positive effects of Echinacea in preventing colds, flu, and upper respiratory tract infections. Other studies show no effect. The use of Echinacea for preventing colds and other ailments is controversial for this reason and because there are also disagreements in the literature about possible carcinogenic effects of Echinacea.

Subjects are then asked

- a. How can scientists disagree about whether or not taking Echinacea prevents colds and flu? *Water fluoridation subjects answer similar questions about that scenario.*
- b. How can scientists disagree about whether or not taking Echinacea causes cancer? *Water fluoridation subjects answer similar questions about that scenario.*

- c. In a controversy like this, is one answer right and one wrong? *If yes*: What would make one right and one wrong? *If no*: Could one be better than the other? What would make it better?
- d. In the case of this controversy, how might scientists go about resolving it? *If they mention experiments*: What might the experiment be in this case? *If no mention of experiments*: Is it possible they might do an experiment? What would an experiment look like in this case?

Smith & Wenk (2006) argue that these three very different sets of questions — about the nature of science, about a statement concerning scientific uncertainty, and about an actual scientific controversy — constitute three different probes of students' epistemologies of science, allowing for confirmation or disconfirmation of the Coherence perspective.

Scoring and results

The three parts of the interview are independently scored using the following three-level scheme based on Carey and Smith (1993):

Table 1

The scheme used by Smith & Wenk to categorize students' epistemologies of science, reprinted from Smith & Wenk (2006).

Table 1

Description of some of the core concepts in three epistemologies of science

Level	Key Differentiations	Nature of Knowledge	Acquisition Processes	Certainty/ Uncertainty
1	No differentiation of ideas and evidence	True beliefs about what happens and what works	Making observations, doing tests, finding answers	Certain knowledge
2	Simple differentiation of ideas and evidence	Well-tested hypotheses Generalizations about how or why something works	Simple hypothesis testing	Transitional (not fully specified)
3	Differentiation among framework theories, specific hypotheses, and evidence	Well-tested coherent theories (explanatory frameworks)	Cycles of hypothesis testing that indirectly test theories	Theory-based uncertainty

Note. Based on Carey and Smith (1993).

The scoring scheme also includes levels 1.5 and 2.5, transitional stages between the three main levels. Using pilot data, Smith & Wenk developed detailed coding schemes (rubrics) for each part of the interview, including indicators and typical statements corresponding to each of the five possible scores (1, 1.5, 2, 2.5, or 3). Each rubric operationalizes the level scheme shown in Table 1; but crucially, each was developed independently. In Part 2, students receive subscores for Awareness of Certainty of Knowledge and How Decide if Disagreement. In Part 3, students receive subscores for Reasons for Controversy and How Resolve Controversy.

Smith & Wenk present data to support their claim that students show epistemological consistency across the three parts of the interview. They are careful to emphasize that this consistency is necessary but not sufficient to support the existence of “intuitive meta-theories,” their instantiation of the Coherence perspective:

One final caveat is in order about the significance of these interrelations for an intuitive metatheories perspective on students’ epistemological stances in science. Within the science education community, there have been debates about the amount of coherence that exists in students’ epistemological thinking between those with knowledge-in-pieces viewpoints versus intuitive theories viewpoints. Those with knowledge-in-pieces views have tended to emphasize the inconsistencies in students’ reasoning across diverse contexts (e.g., Hammer & Elby, 2002; Sandoval & Morrison, 2003) while those in the intuitive-theories camp have emphasized the patterns of coherence (e.g., Smith et al., 2000). Of course, resolving the debate does not simply depend on data about student consistencies or inconsistencies. For example, even if students do have coherent epistemological viewpoints, there could be reasons to expect inconsistency (e.g., knowledge limitations could prevent the application of concepts in certain contexts). Conversely, consistency in student reasoning does not automatically favor the intuitive theories viewpoint (a) in the absence of a detailed conceptual analysis of the contrasting ways that concepts cohere within different theories or (b) if alternative, more compelling reasons for those intercorrelations can be found (e.g., general differences in intelligence, critical thinking, or verbal reasoning ability that affect the pace and ultimate level of development). Thus, we note the reason we think our data argue for some underlying coherence in students’ epistemological thinking is not simply because there were high correlations across epistemological measures but rather because the pattern of observed relations is consistent with our specific conceptual analysis (e.g., consistently differentiating theory and evidence was strongly related to having an understanding of inductive and interpretive uncertainty). Ultimately, we believe that an adequate account of epistemological development will need to integrate the insights from both perspectives. (Smith & Wenk, 2006, p. 775)

Smith & Wenk do not establish epistemological coherence in students’ reasoning

In the next three sections, I show that Smith & Wenk’s article does not support the Coherence perspective. I begin by arguing that their data and analysis are insufficient to establish that individual subjects displayed the same (or approximately the same) level of epistemological sophistication across the three probes.

Claims of Coherence at the expense of Fragmentation, I argue, must satisfy

Evidence of Consistency: In analyses that rank or categorize epistemologies, evidence must show that individual students display roughly the same epistemological level both within and across probes/contexts.

Transparency: Authors must be unusually generous in presenting coded raw data underlying the claim of coherence.

The first condition is straightforward. Still, as diSessa, Gillespie, and Esterly (2004) discuss, theoretical gaps and ambiguities — both within and across theoretical perspectives — preclude clean specification of what counts as “roughly the same” epistemological level, or how much inconsistency is too much. This irreducible imprecision motivates the Transparency condition: to support productive paradigm debate, papers must supply enough coded raw data so that both supporters and skeptics can formulate detailed, evidence-based arguments and counterarguments about the validity of the coding scheme — arguments that could help clarify how the paradigm differences play out in the nitty-gritty details of data analysis.

Smith & Wenk fail to meet these conditions, because (1) High correlation coefficients between students’ epistemological levels on different probes do not establish the epistemological consistency of individual students, (2) Smith & Wenk do not present data showing the degree of intra-student consistency across all three parts of the interview, and (3) they provide insufficient raw and processed data to establish intra-student coherence within each part of the interview. I now present these three critiques in detail.

Irrelevance of correlation coefficients

Discussing the high correlation coefficients between students’ epistemological levels on the three parts of the interview, Smith & Wenk write, “...the intercorrelations among the three aspects of epistemology provide some validation of our measures and of the existence of different coordinated systems of concepts for thinking about science.â To see why correlation coefficients do not contribute to the argument for epistemological coherence, consider Table 2, showing *hypothetical* student data for Smith & Wenk’s experiment. Since every student but Hari ranges over a full epistemological level on a scale of 1 to 3; these results would *not* indicate epistemological coherence. Yet, the Pearson correlations ($r_{12} = .66$, $r_{13} = .81$, $r_{23} = .68$) are as high as Smith & Wenk’s actual correlations. That’s because students who score higher on part 1 tend to score higher on parts 2 and 3, too. The correlation coefficient r_{13} “ignores” the fact that the hypothetical students score low on Part 1 and high on Part 3, indicating lack of epistemological consistency across those two parts.

Table 2

Fake data to illustrate that high correlation coefficients between the three parts of Smith & Wenk's interviews do not indicate intra-student consistency.

	Part 1	Part 2	Part 3
Anna	1	1.5	2
Bob	1.5	2	2.5
Carl	1.25	1.5	2.25
Denise	1.25	1.75	2.25
Ellen	1.25	1.75	2.25
Frank	1	2	2.25
Giselle	1.5	1.75	2.5
Hari	2	2.25	2.5

Just as high correlations do not necessarily indicate coherence, low and negative correlations do not always indicate lack of coherence. In table 3, the correlations are $r_{12} = .12$, $r_{13} = -.19$, and $r_{23} = .25$. Yet, each student ranges over half a level or less. When data “clump” tightly, correlation coefficients are sensitive to small changes in the data.

Table 3:

Fake data to illustrate that low and negative correlation coefficients do not indicate lack of intra-student consistency across the three parts of the interview.

	Part 1	Part 2	Part 3
Alice	2	1.5	2
Bill	1.75	2	2.25
Camisha	2.25	2.5	2
Dennis	2	1.5	2
Edward	2.25	2	1.75
Francisco	2	2.25	2
Gil	1.75	2.25	1.75
Huriko	2	2.5	2.5

In summary, correlation coefficients between students' epistemological levels on different probes do not contribute to arguments for or against students' coherence across those probes. To be fair, correlation coefficients constitute a tiny fraction of the analysis that Smith & Wenk marshal to support their argument for the Coherence perspective. They *do* provide other, more telling evidence. But not sufficient evidence, I now argue, to establish coherence.

Lack of sufficient evidence of consistency across contexts

To establish students' consistency across contexts, Smith & Wenk would need to present the “epistemological range” of each student across the three parts of the interview, i.e., the difference between the student's highest and lowest level, or some

other measure of intra-student epistemological variability. In Table 3 above, for instance, everyone but Francisco has an epistemological range of 0.5 levels. More compact, of course, is a histogram or paragraph summarizing how many students had an epistemological range between 0 and 0.1, between 0.1 and 0.2, and so on.

Smith & Wenk do not present data that allow us to read off or infer the distribution of students' epistemological ranges across the three parts of the interview. For example, the scatter plot in their Figure 1 allows readers to determine, for each subject, the difference between her score on the *second segment* of Part 1 and the *average* of her scores in Parts 2 and 3; but readers cannot inspect the difference between the Part 2 and Part 3 scores. Consider also their Table 12, reproduced here as my Figure 1.

Table 12

Relation between average Nature of Science score in Part 1 and awareness of uncertainty in Part 2

Average Score	Awareness of Uncertainty in Part 2			
	Certain Knowledge Level 1	Some Certain; Some Uncertain Level 1.5	Inductive Uncertainty Level 2	Interpretive Uncertainty Level 2.5
<1.70	8	2	0	0
1.70–1.90	4	4	4	1
> 1.90	0	2	6	4

Figure 1. Smith & Wenk's Table 12, cross-tabulating Part 1 scores and Awareness of Uncertainty subscores from Part 2.

As Smith & Wenk emphasize, students who scored lowest on Part 1 (<1.70) scored below level 2 in the Awareness of Uncertainty segment of Part 2, while students who scored highest in Part 1 (>1.90) almost all scored level 2 or higher in that segment of Part 2. Although these results establish some degree of intra-student epistemological consistency across interview parts (probes),¹ we cannot evaluate exactly how much, not just because Part 3 and some of Part 2 are excluded, but also because of the coarse binning of Part 1 scores. For students who scored low (<1.70) in Part 1, the difference between their Part 1 scores and their Part 2 awareness-of-uncertainty subscores could range from 0 to .7 levels, and we don't know the distribution within that range.

At another point, Smith & Wenk discuss students' epistemological consistency across all three parts of their interviews:

Similarly, there were clear relations between how students reasoned about a specific controversy (in Part 3) and their differentiation of evidence/hypotheses/theories (in Part 1) and their understanding of the uncertainty of scientific knowledge (in Part 2). For example, all students who systematically engaged with the controversy in deeper ways (e.g., by

¹ Smith & Wenk acknowledge the greater variability between levels of the middle group of students, those scoring between 1.70 and 1.90 on Part 1.

showing awareness that the scientists may have made interpretive errors) had been aware of the uncertainty of scientific knowledge in Part 2. Further, all of these students had consistently differentiated ideas and evidence in the Nature of Science Interview (Part 1); they also had at least one and often several Level 2.5 scores on this part as well. (p. 768)

Undeniably, the subgroup of students discussed in this paragraph display some consistency across all three parts. But we cannot infer the degree of consistency — the students' epistemological ranges. Knowing that students answered at least one Part 1 question at level 2.5 doesn't tell us whether their (average) score for that part was 2.5, 2, or lower. And having been "aware of the uncertainty of scientific knowledge in Part 2" refers to a Part 2 *subscore* (of level 2 or 2.5), with no reference to the overall Part 2 score.

In summary, Smith & Wenk do not provide the distribution of epistemological ranges of individual students across all three parts of the interview. Readers therefore cannot evaluate and debate whether the degree of intra-student epistemological consistency support claims of theory-driven (actually, meta-theory-driven) coherence.

Lack of sufficient evidence of consistency within contexts

Smith & Wenk do not always show the epistemological ranges of individual students across the subscores within a given interview part. Their Table 10 (my Table 4 below), however, illustrates a great way to present such data. For the two subscores of Part 3, we see that $4+15+4 = 23$ students had an epistemological range of 0 levels; $2+6+4 = 12$ students had an epistemological range of 0.5 levels; and no students had a higher range. Arguing for coherence, Smith & Wenk would presumably emphasize that two thirds of students were completely consistent within Part 3, and the other third varied over just half a level, which is small. A skeptic might counter that half a level is large, given that Part 3 subscores span only 1 level, from 1.5 to 2.5. My point is that Table 10 provides sufficient detail to support this kind of debate.

Table 4

Smith & Wenk's Table 10, cross-tabulating the two subscores students received in Part 3 of the interview.

Table 10

The relation between student ideas about reasons for controversy and how resolved

Reasons for Controversy	How Controversy Resolved		
	Test More People (Level 1.5)	Do Similar Experiments; Make Simple Comparisons (Level 2)	Do Controlled Experiments; Test in Multiple Ways (Level 2.5)
Not enough data; different people (Level 1.5)	4	2	0
Looked at different variables (Level 2)	6	15	4
Mistaken/conflicting interpretations (Level 2.5)	0	0	4

By contrast, in a similar cross-tabulation of the two subscores from Part 2 (their Table 7), Smith & Wenk bin together level 1 and level 1.5 responses, making it impossible to read off the distribution of epistemological ranges as I did from Table 10.

Addressing the issue of intra-student consistency in Part 1, Smith & Wenk write,

Clearly, there was some variability in students' level of responding from question to question, reflecting in part the degree of scaffolding provided by the question. Nonetheless, by averaging student scores across all 13 questions in the Nature of Science Interview, we can get one index of the consistency with which students were differentiating between theory and evidence throughout the interview. When we did this, we found that 10 students had average scores between 1.38 and 1.67, with most of their responses at Level 1 or Level 1.5. For these students, clearly differentiating theory and evidence was not the norm. In contrast, 12 students had average scores between 1.92 and 2.23, with most of their responses at Level 2 or Level 2.5. These students thus provided strong evidence of consistently differentiating between theory and evidence throughout the interview. Finally, 13 students had average scores between 1.71 and 1.87, and were more evenly split between Level 1.5 and 2 responding. (p. 757)

Smith & Wenk must provide more details to support the claim of coherence. For the 10 bottom-scoring students, "most of their responses" were at level 1 or 1.5. But how many is "most?" The same ambiguity applies to their 12 top-scoring subjects, among whom "most" responses were at level 2 or 2.5. Based on the given information, we cannot tell whether most students' responses ranged over a full level (or more) or clumped into a tighter range.

Even if students received consistent subscores within a given part of the interview, Fragmentation advocates would want to see more of the coded raw data. In a paradigm debate, researchers often dispute the salience of each others' coding schemes, since such schemes often operationalize deeply-embedded assumptions about the underlying construct (see, e.g., Elby & Hammer, 2001). Debates about these disagreements can occur only if participants provide copious coded raw data. To be fair, Smith & Wenk do an unusually thorough job of presenting transcript snippets epitomizing different epistemological levels. Length limitations might have disallowed large chunks of additional transcript. An appendix on the web sidesteps such restrictions. To encourage the kind of deep data sharing needed in a paradigm debate, journals should encourage authors to post as much of their data as possible on the web. My research group has had no trouble obtaining IRB (Human Subjects) approval, and individual subjects' consent, for publishing anonymized transcripts both in print and on the web.

In summary, I have not argued that Smith & Wenk's subjects are epistemologically incoherent. I've argued that we just don't know: Smith & Wenk do not provide sufficient evidence of intra-subject consistency within and across the three parts of their interviews to support claims of coherence. In making this critique, I

illustrated more generally what does and does not count as evidence of coherence within the high standards of accountability demanded by a paradigm debate.

Accounting for coherence in Smith & Wenk's data

I now table the issues of the previous section and assume that Smith & Wenk's subjects displayed coherent epistemological stances. I discuss how to account for that coherence without attributing intuitive meta-theories or other theory-like cognitive structures.

Smith & Wenk acknowledge that epistemological consistency alone does not favor the Coherence perspective:

...consistency in student reasoning does not automatically favor the intuitive-theories viewpoint (a) in the absence of a detailed conceptual analysis of the contrasting ways that concepts cohere within different theories or (b) if alternative, more compelling reasons for those intercorrelations can be found (e.g., general differences in intelligence, critical thinking, or verbal reasoning ability that affect the pace and ultimate level of development). Thus, we note the reason we think our data argue for some underlying coherence in students' epistemological thinking is not simply because there were high correlations across epistemological measures but rather because the pattern of observed relations is consistent with our specific conceptual analysis (e.g., consistently differentiating theory and evidence was strongly related to having an understanding of inductive and interpretive uncertainty). (Smith & Wenk, 2006, p. 775)

Notice that Smith & Wenk's examples of alternative explanations of students' epistemological consistency — “general differences in intelligence, critical thinking, or verbal reasoning ability” — are general cognitive abilities of the student, presumably corresponding to robust cognitive structures activated across a wide range of contexts, including an interview probing epistemology. Smith & Wenk don't rule out other alternatives, of course; but the tenor of these examples is that if students' epistemological consistency isn't explained by one robust cognitive structure (an intuitive meta-theory), then it's explained by *other* robust cognitive structures.

By contrast, my explanation of epistemological consistency in Smith & Wenk's subjects centers around contextual factors: *priming* (Hirshman, Snodgrass, Mindes, & Heenan, 1990; Marsh, Bink, & Hicks, 1999), the idea that activating a knowledge element or pattern of thought during a task can influence how the subject thinks during a follow-up task; and *contextual stability*, the idea that a pattern of thought, once established and reinforced, is likely to persist if the context stays comparatively stable (Luchins, 1942), as is the case in Smith & Wenk's experiment (I'll argue). My discussions of priming and contextual stability are largely phenomenological, compatible with both the Coherence and Fragmentation perspectives.

Priming

I'll introduce my argument with an analogy. Suppose a nutrition education researcher investigates students' views about fats. During part 2 of the interview, the researcher prompts subjects to comment upon a (fictional) student's assertion, "Some fats are good and you can eat as much of them as you want, while others are so bad they should always be avoided." The subject spends 10-15 minutes responding to questions about that assertion. Then, in part 3, the interviewer unveils three boxes of snack crackers, whose nutrition labels indicate different mixes of saturated, unsaturated, and trans fats. The subject chooses a box to take home and explains his choice.

Suppose subjects' beliefs about fats expressed in part 2 mesh with the beliefs they enact and explain in part 3. Would that count as evidence of coherence in students' professed and enacted beliefs about fats? No: After discussing their beliefs about fats for 10-15 minutes, subjects have those thoughts in their heads when presented with the snack choice. In other words, discussing their beliefs in part 2 *primes* subjects to act upon and reiterate those beliefs when choosing crackers immediately afterward in part 3, even if those beliefs aren't stable or robust. For this reason, consistency across parts 2 and 3 is not good evidence that a theory-like cognitive structure explains the consistency.

These considerations also apply to Parts 2 and 3 of Smith & Wenk's experiment. In Part 2, subjects hear a hypothetical student's statement about uncertainty in science, and they answer questions about the nature of this uncertainty and how to resolve it. Immediately afterwards, in Part 3, students hear a real-life scientific controversy and answer questions about the reasons for the disagreement and how it could be resolved. Surely, subjects' thoughts about scientific uncertainty and resolution activated in Part 2 could influence their responses to the scientific controversy in Part 3. This priming could help to account for consistency between Part 2 and Part 3 responses.

Priming can also partly explain consistency between Part 1 and Part 2 responses. The final questions in Part 1 are "Do scientists ever change their hypotheses or theories? When would they do that and why?" and "Do you think it is easier to change a hypothesis or a theory? Why?" These questions activate thoughts about why scientists and perhaps other people change their minds, which relates closely to Part 2's focus on the nature and resolution of disagreements in science. So, the thoughts activated at the end of Part 1 could influence subjects' thinking in Part 2.

In summary, priming can at least partly explain why students display epistemological consistency across the three parts of Smith & Wenk's interviews, even if students do not possess intuitive meta-theories.

Contextual stability

Even without invoking detailed theoretical considerations, we know that a pattern of thought can persist coherently within a context even when the pattern doesn't necessarily arise from a stage- or theory-like cognitive structure. For example, "einstellung" experiments going back to the 1940s establish that a subject's approach to

solving a particular kind of problem, once established and practiced, often persists even when no longer appropriate (Luchins, 1942). Therefore, experiments like Smith & Wenk's can favor Coherence over Fragmentation only by probing students across different contexts. Smith & Wenk may be acknowledging this point in emphasizing that the three parts of their interview are "three different types of probes" (p. 752). I now argue, however, that the three parts of the interview do not constitute different contexts.

Because "context" is theoretically ambiguous both within and especially across the Coherence and Fragmentation perspectives (diSessa et al., 2004), I cannot make this argument watertight. And Smith & Wenk emphasize, correctly, that answering brief questions about the nature of science (Part 1) differs from responding to an assertion about the certainty and uncertainty of scientific knowledge (Part 2), which differs from analyzing a scientific controversy (Part 3). So, the *fine-grained subject matter* and *task structure* components of the context both change. But other aspects of the context remain stable over the entire interview, including

- The *physical setting*;
- The *social setting*, who is talking to whom, and the relationship between them;
- The *purpose* of the interaction — whatever the subject perceives that purpose to be, her perception has no reason to shift;
- The *general subject matter*, the nature of scientific knowledge and uncertainty;
- The *general task structure*, answering specific questions posed by the interviewer.

The stability of these contextual factors makes it plausible that a pattern of epistemological thought, once established, could persist over the three parts of the interview — especially given the priming discussed above.

In summary, priming and contextual stability can partly account for subjects' epistemological coherence across the three parts of Smith & Wenk's interviews, with no need to attribute theory-like cognitive structures such as intuitive meta-theories. Therefore, Smith & Wenk's experimental results cannot favor the Coherence perspective over the Fragmentation perspective.

The Fragmentation perspective can account for local coherences

This section shows how the Fragmentation perspective can account for students' epistemological coherence in Smith & Wenk's experiment.

Sometimes, the Fragmentation perspective is taken to imply complete fragmentation and incoherence in students' thinking. That is not the view of Fragmentation advocates (diSessa, 1996; Elby & Hammer, 2010; Hammer, Elby, Scherr, & Redish, 2005). They expect local epistemological coherences, including belief-like consistencies and stabilities, within a given context (diSessa, Elby, & Hammer, 2002;

Hammer, 1994) and sometimes across contexts (Lising & Elby, 2005). What distinguishes them from Coherence advocates, at the phenomenological level, is the expectation that a student has *multiple* epistemological stances available to her, and that contextual cues and other factors can greatly influence which locally-coherent epistemological stance turns on in a given scenario.

To flesh out these points, I start by briefly summarizing how the “epistemological resources” framework for describing student epistemologies can account, in general, for students’ epistemological coherence in a given context (Hammer & Elby, 2002; Hammer et al., 2005). According to this framework, students’ epistemologies are fragmented in the sense that different combinations of epistemological resources — our name for the units or “atoms” of epistemic thought — get triggered by different circumstances (Hammer & Elby, 2002; Hammer et al., 2005; Louca, Elby, Hammer, & Kagey, 2004). In some cases, a network of activated resources is unstable and turns off quickly. In other cases, the network of resources is *locally coherent*, i.e., the activations of the individual resources are mutually reinforced by each other and/or by contextual cues and other features of the environment, leading to stability (Hammer et al., 2005). The stable activation of a locally coherent network of resources — an *epistemological frame* — corresponds to a coherent epistemological stance observed in the subject (Elby & Hammer, 2010; Hammer et al., 2005; Redish, 2004). In epistemological experts, particular epistemological frames have become compiled into full-fledged epistemological beliefs or theories. By contrast, in novices, an epistemological frame may depend on contextual factors for its stability, and different contexts can trigger different frames. So, the coherent epistemological stances displayed by some of Smith & Wenk’s subjects might be local coherences rather than global coherences resulting from an intuitive meta-theory. That is why epistemological coherence in their subjects does not favor Coherence over Fragmentation.

Indeed, previous work (Elby & Hammer, 2010; Rosenberg et al., 2006) introduced a specific locally-coherent network of resources — a particular epistemological frame — that can account for a central pattern of epistemological coherence observed by Smith & Wenk. Discussing correlations in their data, Smith & Wenk write,

In particular, making a consistent differentiation between ideas and evidence was associated with having an appreciation of the general uncertainty of scientific knowledge, with having an appreciation of the deeper interpretive issues in scientific controversies, and with understanding some of the domain-specific methods scientists have developed for resolving these controversies. (p. 774)

An epistemological frame called *Causal Storytelling* can account for that coherence. I lack space here to lay out the particular epistemological resources involved in this frame and why they are mutually reinforcing; see Elby & Hammer (2010) and Rosenberg et al. (2006) for details. Very briefly, *Causal Storytelling* combines the view that knowledge is built out of “raw materials,” the expectation that pieces of knowledge fit into a narrative, the expectation that those pieces should be causally linked (rather than

disconnected), and other units of epistemic thought. The resulting frame corresponds to the epistemological stance that knowledge generation consists of constructing causal stories. With this frame activated, the subject distinguishes evidence (pieces of information) from ideas (the story into which those pieces fit, and the causal connections between them). That frame also allows the subject to see the possibility that scientific uncertainty arises not just for simplistic reasons (“one of the scientists made wrong observations”) but also for deeper interpretive reasons, specifically, constructing different causal stories (mechanistic explanations) out of the same evidence. In brief, the locally coherent activation of *Causal Storytelling* can account for Smith & Wenk’s result that subjects who consistently distinguished ideas from evidence also tended to view uncertainty in sophisticated ways. And earlier work establishes that *Causal Storytelling* is not always a context-independent belief, but rather, an epistemological frame that turns on in response to some cues but not in response to others (Elby & Hammer, 2010; Rosenberg et al., 2006).

Testing Coherence versus Fragmentation

What *would* constitute a fair test of Coherence vs. Fragmentation in the realm of personal epistemologies? Theoretical complexities and ambiguities about the notions of coherence and context preclude a clean answer (diSessa et al., 2004). Still, I now propose a variant of Smith & Wenk’s experiment that both camps might accept as a fair wedge experiment. It consists of two clinical interviews, administered in different weeks in different places by different interviewers, so that Part A doesn’t prime Part B.

Part A: Primed response to assertion about uncertainty

My Part A is a variant of Smith & Wenk’s Part 2 with elements of Part 3 mixed in. As in Smith & Wenk’s Part 2, students respond to this assertion from a hypothetical student:

Everyday, in more and more areas of science the right answer is known. In areas where the right answer is known, I look to experts to tell me what is right. In areas where no right answer is known, I think anyone’s opinion is as good as another’s.

But first, students read about a controversy over the cause of AIDS:

Although most scientists think AIDS is caused by a virus called HIV, a few scientists, including an eminent researcher who helped to discover the first cancer gene, think AIDS is caused by something else. Several years ago, the president of a country suffering from an AIDS epidemic became largely convinced by scientists who dispute the link between HIV and AIDS. According to health-care advocates, his doubts may have contributed to several years of inadequate attention to prevention of HIV transmission and inadequate resources spent on anti-viral medicines for AIDS sufferers in his country. Most health-care advocates also emphasize that the country’s poverty and history of

colonial oppression, and the unwillingness of Western pharmaceutical companies to supply cheaper AIDS medications, is a deeper cause of the AIDS crisis there.

After reading and hearing this paragraph and the hypothetical student's assertion, the subject answers variants of Smith & Wenk's questions designed to keep the AIDS controversy in mind. For example, instead of "Do you agree with this statement? Why or why not?" the subject is asked "Do you agree with this statement? Why or why not, in light of the AIDS controversy?"

I hypothesize that the AIDS controversy will disproportionately trigger a frame called *Fact Accumulation* (Elby & Hammer, 2010; Hammer et al., 2005; Redish, 2004; Rosenberg et al., 2006), the stance that knowledge generation is mostly a matter of accumulating information. This frame supports the idea that there's a right answer about whether HIV causes AIDS and that scientific disagreement arises from mistakes rather than deep interpretive issues. So, I predict more level 1 and level 1.5 responses, and fewer level 2.5 responses, than Smith & Wenk obtained in their Part 2.

Part B: Resolving arguments

Ideally, these interviews occur in a non-academic setting such as a quiet café, with an informally-dressed interviewer, different from the Part A interviewer. The interviewer (and all recruitment material!) presents the purpose as seeing how different people evaluate and try to settle substantive, friendly arguments. The subject reads the following:

A friendly argument between two weight-loss specialists

Two nutritionists, Danielle and Rani, went to college together and are close friends. Both devote their careers to helping people lose weight. They have different opinions, however, about the most effective way to help clients. Danielle says the key is sleep:

Danielle: For the last several years, I've asked all of my clients to keep a sleep log. The ones who consistently got 8 or more hours of sleep lost weight more quickly than the others. And this makes sense, because inadequate sleep changes your metabolism, making you burn energy less efficiently.

Rani, by contrast, thinks the most effective treatment is dietary supplements that increase metabolism, making the body burn energy more quickly and reducing feelings of tiredness, which might help clients exercise more.

Rani: After working with a doctor to check my clients' medical conditions to make sure the pills are safe for them, I had half of my clients who were interested take these supplements in my presence, so I could be sure they didn't take too many. And they lost weight faster than the others.

Danielle: I tried the same thing one time, working with clients who came to an evening clinic I run with a doctor. But the patients who took supplements didn't lose weight any faster than my other clients. So I'm not really convinced the supplements help.

After reading and re-reading this passage, the subject responds to variants of Smith & Wenk's prompts from their Part 3:

- a. How can those two friends disagree about whether the supplements help people lose weight?
- b. How can they disagree about the importance of sleep?
- c. In a disagreement like this, is one answer right and one wrong? *If yes:* What would make one right and one wrong? *If no:* Could one be better than the other? What would make it better?
- d. In the case of this argument, how might the friends go about resolving it? *If they mention experiments:* What might the experiment be in this case? *If no mention of experiments:* Is it possible they might do an experiment? What would an experiment look like in this case?

The non-academic atmosphere, the framing of the disagreement as a friendly argument rather than a scientific controversy (at least until part d), and the familiarity of the content are designed to trigger everyday resources people have for argumentation, which I hypothesize are closely linked to intuitive sense-making. ("School science," by contrast, may systematically suppress many students' argumentation and sense-making (Lemke, 1990).) The activation of *argumentation* resources and students' sense of mechanism triggers and helps to stabilize *Causal Storytelling*, leading to level 2 and level 2.5 responses. Specifically, I predict that many subjects will propose controlled experiment(s) attending to both supplements and sleep, and bringing theoretical considerations of mechanism into the discussion (hints of level 3!).

In summary, I have outlined an experiment for which resources-based considerations lead to the prediction that students will show greater epistemological sophistication on Part B than on Part A. By contrast, if students' answers are driven by coherent stages, theories, or meta-theories, we have no reason to expect differences in epistemological sophistication between the two parts.

Is that a fair test?

A theory/stages advocate could share my intuition that students will do better on Part B while denying that this experiment is a fair test of Coherence vs. Fragmentation. After all, Part B provides more scaffolding, presenting both sides of an argument and even mentioning a specific causal mechanism. As Smith & Wenk write,

At present, although there is much data that students can look more or less sophisticated in different contexts depending upon the amount of scaffolding supplied (Bell & Linn, 2002), it is not always clear that the

strategies students use in those situations depend on the students' own epistemological conceptualizations of the problem. (p. 752)

And indeed, if the scaffolding is explicitly epistemological or very directive, then a theory/stages advocate can deny that the scaffolded performance reflects the subject's epistemology. In my proposed experiment, by contrast, the scaffolding in part B and the "reverse scaffolding" in part A are not explicitly epistemological. My scaffolding doesn't tell students how to view or treat knowledge. Instead, the scaffolding in Part B is conceptual, using a familiar conceptual terrain with plausible glimmers of mechanism, and contextual, striving to frame the scenario as informal argumentation rather than scientific debate. If conceptual and contextual cues can "knock" students into different epistemological stances in predictable ways, then attributing to students a stable, robust, context-independent meta-theory or theory provides less predictive and explanatory power than ascribing epistemological resources that come together into different locally coherent networks (frames) in response to different cues.

Conclusion

In this response to Smith & Wenk (2006), I tried to further the debate between the Coherence and Fragmentation perspectives in the personal epistemology research community. Arguing that Smith & Wenk do not meet the high burden of proof and transparency needed in a paradigm debate, I highlighted more generally what kind of data presentation would support claims of coherence in a way that furthers the debate. However, even a high degree of epistemological consistency across the three parts of Smith & Wenk's interview would not favor Coherence over Fragmentation. Priming and contextual stability can account for that coherence; and a specific resources-based model can account for a particular epistemological coherence underscored by Smith & Wenk. I hope that my proposed variant of Smith & Wenk's protocol will provoke further debate about these important paradigmatic issues.

Acknowledgment

I thank David Hammer for comments and discussion. This work is supported by the National Science Foundation under grants REC-0440113 and ESI-0455711. Any opinions, findings and conclusions or recommendations expressed are those of the author and do not necessarily reflect the views of the Foundation.

References

- Anderson, T., Tolmie, A., Howe, C., Mayes, T., & Mackenzie, M. (1992). Mental models of motion. In Y. Rogers, A. Rutherford & P. A. Bibby (Eds.), *Models in the mind: Theory, perspective, and application* (pp. 57-71). New York: Academic Press.
- Carey, S., Evans, R., Honda, M., Jay, E., & 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. In R. Driver (Ed.), *Students' conceptions and the learning of science, a special issue of International Journal of Science Education*. (Vol. 11, pp. 514-529).
- Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, 28(3), 235-252.
- Clark, D. B. (2003). Analyzing student knowledge integration: Theories or pieces? Paper presented at the National Association of Research in Science Teaching Conference, Philadelphia.
- diSessa, A. A. (1996). What do "just plain folk" know about physics? In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching, and schooling* (pp. 709-730). Oxford: Blackwell Publishers.
- diSessa, A. A., Elby, A., & Hammer, D. M. (2002). J's epistemological stance and strategies. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 237-290). Mahwah, NJ: Lawrence Erlbaum.
- diSessa, A. A., Gillespie, N., & Esterly, J. B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, 28(6), 843-900.
- Elby, A. (2000). What students' learning of representations tells us about constructivism. *Journal of Mathematical Behavior*, 19, 481-502.
- Elby, A., & Hammer, D. M. (2001). On the substance of a sophisticated epistemology. *Science Education*, 85, 554-567.
- Elby, A., & Hammer, D. M. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice* (409-434). Cambridge, UK: Cambridge University Press.
- Hammer, D. M. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12(2), 151-183.

- Hammer, D. M., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169-190). Mahwah, NJ: Lawrence Erlbaum.
- Hammer, D. M., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89-120). Greenwich, CT: Information Age Publishing.
- Hirshman, E., Snodgrass, J. G., Mindes, J., & Heenan, K. (1990). Conceptual Priming in Fragment Completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 634-647.
- Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during Online searching. *Educational Psychologist*, 39(1), 43-55.
- 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(1), 88-140.
- Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*. Mahwah, N.J.: Lawrence Erlbaum.
- Ioannides, C., & Vosniadou, S. (2002). The changing meanings of force. *Cognitive Science Quarterly*, 2, 5-61.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment : understanding and promoting intellectual growth and critical thinking in adolescents and adults* (1st ed.). San Francisco: Jossey-Bass Publishers.
- Lemke, J. L. (1990). *Talking Science: Language, Learning & Values*. Norwood NJ: Ablex.
- Lising, L., & Elby, A. (2005). The impact of epistemology on learning: A case study in introductory physics. *American Journal of Physics*, 73(4), 372-382.
- Louca, L., Elby, A., Hammer, D. M., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, 39(1), 57-68.
- Luchins, A. S. (1942). Mechanization in problem solving: the effect of Einstellung. 54 (whole number 248).
- Marsh, R. L., Bink, M. L., & Hicks, J. L. (1999). Conceptual priming in a generative problem-solving task. *Memory and Cognition*, 27(2), 355-363.
- Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. In E. F. Redish, C. Tarsitani & M. Vincentini (Eds.),

- Proceedings of the Enrico Fermi Summer School, Course CLVI* (pp. 1-63). Bologna: Italian Physical Society.
- Rosenberg, S. A., Hammer, D., & Phelan, J. (2006). Multiple epistemological coherences in an eighth-grade discussion of the rock cycle. *Journal of the Learning Sciences, 15*(2), 261-292.
- Samarapungavan, A., & Wiers, R. W. (1997). Children's thoughts on the origin of species: A study of explanatory coherence. *Cognitive Science, 21*(2), 147-177.
- 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*(4), 369-392.
- Schommer, M. (1990). The effects of beliefs about the nature of knowledge in comprehension. *Journal of Educational Psychology, 82*(3), 498 - 504.
- 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*(3), 349-422.
- Smith, C. L., & Wenk, L. (2006). Relations among Three Aspects of First-Year College Students' Epistemologies of Science. *Journal of Research in Science Teaching, 43*(8), 747-785.