Explanations that Make Sense: Accounting for Students' Internal Evaluations of Explanations

Shulamit Kapon, Faculty of Education, University of Haifa, 199 Aba Khoushy Ave., Mount Carmel, Haifa 3498838, Israel, shulamit.kapon@edu.haifa.ac.il
Orit Parnafes, Education Department, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 8410501, Israel, Oritpa@bgu.ac.il

Abstract: This paper discusses students' internal evaluation of explanations, as opposed to external evaluations of explanations that are based on compatibility with formal scientific explanations. We propose a theoretical construct, metaphorically termed *the internal metric for certainty in explanation,* to account for the process of internal evaluation. We suggest that the internal metric is multidimensional, and discuss three dimensions: (1) intuition; (2) local coherence; (3) mechanistic reasoning. We argue that these three dimensions are highly active in self-evaluating explanations about phenomena in the physical world. We operationalize these dimensions into empirically traceable terms. We illustrate our framework through an analysis of an episode of a student's reasoning about what causes a plastic bottle to shrink when air is pumped out of it. The analysis demonstrates that the framework can explain conviction in an explanation, as well the preference for one explanation over another.

Introduction

One of the most common words in young children's vocabulary is "why." This word, indicating curiosity, points to children's essential need to understand and explain the world around them, a need that does not disappear as they grow older. From a very young age, humans try to explain the phenomena they encounter (Gopnik, 1998) for sense making purposes (Berk, 1985; Chi, Bassok, Lewis, Reimann, & Glaser, 1989) and for social purposes (Dagher & Cossman, 1992). Moreover, humans frequently evaluate various explanations of the same phenomenon, choosing the one that seems better to them and rejecting the one that seems worse (Harman, 1965; Kapon & diSessa, 2012; Keil, 2006). The assumption that guides this work is that the ability to self-evaluate explanations according to internal and not only external measures, is important for sense making and in the progression from one explanation to another. Although students' explanations are commonly evaluated by means of their compatibility to formal socially-constructed scientific explanations, students' internal evaluations of explanations may have different dimensions, and may rely on different resources. This study aims to explore the dimensions by which students evaluate their own explanations and those of others.

We suggest a theoretical construct that is aimed to account for the process of internal evaluation, which we metaphorically term *the internal metric for certainty in explanation*. Our goals in this paper are two-fold. First, we argue based on the literature that humans indeed possess such internal metric, and that this metric is highly contextually dependent. Second, we aim to operationalize this elusive metaphorical term into empirically traceable terms for the special case of explanations of phenomena in the physical world. We illustrate our framework through an analysis of an episode of a student's reasoning about what causes a plastic bottle to shrink when air is pumped out of it.

What Counts as a Good Explanation: Extrinsic and Intrinsic Considerations

The word metric implies standard measurement. Philosophers of science have long been debating the nature, characteristic features and developmental patterns of formal scientific explanations (e.g., Bechtel & Abrahamsen, 2005; Friedman, 1974; Machamer, Darden, & Craver, 2000; Salmon, 1989; Toulmin, 1972). Indeed, the most commonly used metric in educational practice to evaluate an explanation of a phenomenon in the physical world is the degree of its adherence to the formal scientific explanation, which comes to bear for instance as regards the nature of the argument students generate (e.g., Driver, Newton, & Osborne, 2000), and the way scholars conceptualize students' learning progressions (e.g., Mohan, Chen, & Anderson, 2009). The degree of adherence to the formal scientific narrative is an external metric that is culturally shaped by the nature of the discipline. In contrast, the metric that concerns us is an internal one that functions at the individual sense making level.

The basic idea that humans have an intrinsic sense of what counts as a good explanation is not new. Harman (1965), for instance, discussed people's inference of the best explanation, namely, the ability to draw inferences that do not follow logically from premises, but are rather based on temporal assumptions about what would be the best explanation. Keil, Rozenblit, and Mills (2004) noted the incomplete nature of understanding, terming it the "illusion of understanding", and suggested that asking people to explain makes people aware of gaps in their understanding. Gentner and colleagues pointed out that people evaluate analogical explanations based on their structural similarities (Gentner & Gentner, 1983), and Holyoak and Thagard (1995) emphasized the role of pragmatic constraints (e.g., the goal of the explanation) in this evaluation. diSessa (1993) suggested that as humans interact with the physical world they develop simple unitary and non verbal schemas (phenomenological primitives, or p-prims) that when activated account for people's comfort with certain situations or surprise in others. Hence, explanations that resonate with an activated p-prim will be evaluated as more plausible (diSessa, 1993, forthcoming; Kapon & diSessa, 2012). Moreover, diSessa (1993) suggested that each p-prim has a *reliability priority* which results from processing feedback that can reinforce or undo the initial activation. Reliability priority expresses the degree of conviction a person attaches to a particular p-prim in a particular context of thought. High reliability priority means that it is unlikely that the p-prim will be deactivated by subsequent processing. Kapon and diSessa (2012) noted that often people activate conflicting explanatory primitives (not necessarily just p-prims) and they internally negotiate their relative priority. Parnafes (2012) conceptualized explanations as comprised of carefully assembled knowledge elements of different types. She argued that people are sensitive to the local coherence of these elements; namely, knowledge elements that comprise an explanation that makes sense cannot suggest conflicting inferences.

The above review suggests that people have an intrinsic sense of what counts as a good explanation, and thus the assumption that they have developed an internal metric for this purpose is reasonable. The question that concerns us is whether we can operationalize this internal metric into external traceable terms in sufficient detail that will allow us to explain the relative evaluations of competing explanations that individuals constantly make (i.e., why does a person prefer one explanation over another) by different values along this metric.

The Internal Metric for Certainty in Explanations: Three Dimensions

We believe that the internal metric for certainty in explanations are multidimensional, and that different dimensions are active in different contexts of thought. This is not a radical assumption. In a review entitled *Explanation and Understanding* Keil (2006) argued that "explanations can be contrasted in terms of domains roughly corresponding to academic disciplines" (p. 232). In fact, it has long been argued that even young children are sensitive to different domains of thought such as naive psychology (human behavior), naive biology (living things), and naive physics (the physical world around us) (Carey, 1985; diSessa, 1993; Inagaki & Hatano, 2002; Keil, 1994). Hence, it is likely that examining the internal metric for evaluating explanations in different domains of thought will help reveal different dimensions of this metric. Here we limit our discussion to a particular domain of thought, explanations about phenomena in the physical world; namely, phenomena that could, at least in principle, also be explained by formal scientific explanations. In this section we operationalize three dimensions of the internal metric that in our view are highly active in self-evaluating explanations about phenomena in the physical world: (1) Intuition; (2) Local coherence; (3) Mechanistic reasoning. We emphasize that these are *not* an exhaustive list of dimensions of the internal metric of certainty in explanations, and there are probably other dimensions. In what follows, we describe each of the three dimensions.

Intuition

Drawing and elaborating on diSessa's (1993) model of p-prims, Kapon and diSessa (2012) developed a model of explanation and change in explanation focusing on knowledge elements (KE) that provide a sense of satisfaction to those judging the explanation. They termed these KE explanatory primitives (e-prims). Like p-prims, e-prims are self-explanatory, unquestioned units of explanation, which students take as simply "the way things are," within the time span of evaluating an explanation. P-prims are a subcategory of e-prims, hence, not all e-prims are p-prims. For instance, a p-prim, by definition, is abstracted from experience in the physical world, whereas the source of an e-prim can be social interactions, language, or even explicit instruction. According to this model, explanations are accepted or rejected on the basis of the individual's convictions concerning particular e-prims and how these primitives fit the context of thought. The e-prims model suggests one dimension to our internal metric of certainty in explanations; namely, that at the base of any conviction lies an e-prim. Table 1 present a list of criteria for identifying e-prims.

Table 1: Criteria for identifying e-prims (Kapon & diSessa, 2012)

Criterion	Definition
Functionality	The knowledge element is explanatorily useful to the goal of reasoning and responsive to the context in which this reasoning takes place.
Obviousness	The knowledge element is regarded with explicit statements or unelaborated confidence and acceptance.
Developmental history	We are able to identify familiar experiences from which the knowledge element could have been abstracted.
Triangulation of expression	The knowledge element reappears frequently in a variety of manifestations during the reasoning process.

Criterion	Definition
Triangulation of form and content	The knowledge element matches documented p-prims or other documented intuitive notions. This criterion cannot always be satisfied, but when it does, we can consider our interpretation to be on a safer ground.

Local Coherence

The idea of consistency or coherence was used by Thagard (2007) for judging the quality of scientific explanations. In reference to Thagard's coherence, Hammer, Elby, Scherr, & Redish (2005) coined the term "local coherence", to highlight that the activation of knowledge elements and the formation of specific coherences are context sensitive. Sherin, Krakowsky, & Lee (2012) also supported this notion by suggesting that when a student explains a phenomenon, a temporary conceptual structure underlying the explanation exhibits consistency which may shift to a different explanation as the activity unfolds. Drawing on the notion of local coherence, Parnafes (2012) suggested that an explanation that makes sense to a student ties together a set of knowledge elements of different types that cohere, locally, in a specific context. A student's (temporary) explanation is viewed as a collection of activated knowledge elements that the student feels fit together at the particular time and for the particular purpose, hence the explanation make sense. Parnafes suggested that a student who is trying to understand a phenomenon goes through iterations of self-explanations, with temporarily stable stages of satisfaction with the explanation she generates in between. This comfort is interpreted as a temporary plateau of local coherence.

Mechanistic Reasoning

Causality plays a privileged role in explanations (Keil, 2003). Moreover, the ability to construct a mechanistic explanation for a physical phenomenon seems to be a developmental achievement (Metz, 1991). We hypothesize that the internal metric of certainty in explanations has a scale of causal strength of an explanation, and that this dimension is active in assessing explanations and self explanations about phenomena in the physical world. Russ, Scherr, Hammer, and Mikeska (2008) adapted an account of mechanism from the philosophy of science (Machamer et al., 2000) and developed a coding scheme that attempts to identify mechanistic reasoning in students' discourse in inquiry-based science classrooms. An underlying assumption of their coding scheme is that evidence of mechanistic reasoning in students' talk does not guarantee that the explanations they generate are scientifically correct. This assumption makes Russ et al.'s framework a promising tool to empirically assess the level of mechanistic reasoning in students' explanations, and whether it correlates with students' relative convictions in these explanations. Table 2 presents the coding scheme in Russ et al.. Note that Russ et al. conceptualized the first seven codes in Table 1 as hierarchically ordered, namely that code #7, chaining, is the highest evidence for mechanistic reasoning. The last two codes according to the Russ et al. framework are not necessarily connected to a particular level of mechanistic reasoning and the numbering of A & B are arbitrary.

Table 2: Coding scheme for mechanistic reasoning in students' talk (Russ et al., 2008)

Code	Definition
(1) Describing the Target Phenomenon	A clear statement or demonstration of the particular phenomenon or result that the student is trying to explain.
(2) Identifying Setup Conditions	Descriptions of the spatial and temporal organization of entities that enable the mechanism to run and produce the phenomenon.
(3) Identifying Entities	Recognizing the objects that affect the outcome of the phenomenon.
(4) Identifying Activities	Identifying the various doings in which the entities engage, articulate the actions and interactions that occur among entities, describe the things that entities do that cause changes in the surrounding entities.
(5) Identifying Properties of Entities	Identifying and articulating general properties of entities that are necessary for the particular mechanism to run.
(6) Identifying Organization of Entities	Identifying how the entities are spatially organized, where they are located, and how they are structured.
(7) Chaining Backward and Forward	Using knowledge about the causal structure of the world to make claims about what must have happened previously to bring about the current state of things (backward) or what will happen next given that certain entities or activities are present now (forward).
(A) Analogies	Using analogies to similar mechanisms in other contexts or fields as a framework for understanding new situations, comparing the target phenomenon to another phenomenon.

Code	Definition
(B) Animations	Using external animated models (gestures, body movements, etc.) to illustrate how certain entities act in the mechanism.

We now turn to a case study in which we highlight two explanations put forward by one student and analyze the internal metric in reference to these two explanations, using our framework.

Illustrative Case Study

The episode that will be analyzed here is drawn from a large corpus of data that was collected as part of the second author's project on the role of self-generated representations for promoting conceptual understanding. As part of the project, the second author and her research group interviewed pairs of students who were asked to generate, discuss and come to an agreement on explanations for phenomena in the physical world while generating representations and elaborating them. Interventions were kept to a minimum during students' discussions, and if made, were done for the purpose of clarifying meanings, or cuing students to consider overlooked conceptual issues. The episode (~ 3.5 minutes) is drawn from the first part of a 90 minute session. The students, two 6th graders, Reut and Natalie (pseudonyms), were trying to explain why a plastic bottle shrinks when air is pumped out of it. The analysis will focus on Reut's reasoning.

We chose the episode and the focus on Reut's reasoning since the explanations that Reut generated at the beginning of the episode and at the end of the episode were very different, and although she seemed quite sure of her explanation at the beginning of the episode she ended with a different one, allowing us to use our model to compare and contrast the two self-generated explanations, and account for Reut's internal evaluation. In each episode, we code and analyze the three dimensions, and then integrate the findings into one integrated metric. The three dimensions were analyzed as follows: (1) intuition: we listed all the knowledge elements in the explanation and those that were likely to be an e-prim, were checked using the criteria listed in Table 1. (2) Local coherence: we reexamined the knowledge elements that were coded for each explanation and examined the degree of fit. (3) Mechanistic reasoning was examined using the codes in Table 2.

Due to space limitations we cannot reproduce the full transcripts. To illustrate the function of the internal metric of certainty in explanation, we focus only on Reut's explanations and often regard the interaction with Natalie and the interviewer as part of the context in which Reut's explanation acquire its meaning, and add them as comments to Reut's explanation. We start by presenting the explanations that Reut generated, and then move to their analysis using our framework to explain why Reut preferred the second over the first. We systematically number each student's utterances to allow us to refer to specific utterances later in our analysis. Utterances are numbered as follows: student name, #of explanation, and a letter that numbers the utterances of this student. Thus, "Reut1B" denotes Reut's 1st explanation, second utterance. Italicized square brackets "*[]*" denote our interpretive and other comments.

The aim of the analysis is to illustrate the operationalization of the internal metric of certainly in explanations. Hence, we examine the two explanations that Reut generated using the metric, explain her initial conviction in the first explanation, and the reason why she later preferred the second explanation over the first via differences between the two explanations along the three dimensions specified above.

Explanation 1

The students watched an experiment in which a plastic bottle shrinks when air is pumped out of it, and they then generated drawings to explain what happened. Using their own drawings they explained the phenomenon to each other, and occasionally the interviewers joined in. Reut and Natalie generated different explanations at the beginning of the episode. Table 3 presents Reut's first explanation.

Table 3: Reut's first explanation (three turns)

Reut1a:	It *[the bottle]* shrank in order to fill the vacuum, to reduce it.
Reut1b:	*[Answering a question from the interviewer "Why should the bottle care about the vacuum?"]* What… *[points to her drawing]* I don't know is why this happens, but I know it does… I even don't know why I know it, but… if there is a vessel that has vacuum inside it and there are particles outside*[Natalie introduced "particles" to the discussion]*, or something… it always enters the vacuum to fill it. I don't know why it happens… hmmm…
Reut1c:	*Points again on her drawing]* this is what I know, that always, if there is a vacuum so something will fill it. This is why here, there is more than – there is vacuum *[points to Natalie's drawing]* and the bottle shrank to fill it out *[demonstrates the change in the shape of the bottle with her hands]*. It's like, shrink inside.

Reut's Initial Evaluation to Natalie's Explanation

Natalie constructed a very different explanation. We provide here only what she said after Reut's turn 1C, since Reut later refers to this explanation.

> Natalie 1D: Because the outside, it's like the pressure on the outside is bigger than the inside. A lot of pressure is activated towards the inside because the...ah... the number of particles outside, say for this area, is bigger than the number of particles inside for this area. So the outside pushes the bottle, to... shrink.

At this point the interviewer asked Reut: "What do you think Reut about what Natalie just said (Natalie 1D)?" Reut replied: "I have no idea if its right or wrong". In other words, at this point she did not think this explanation was better than her own.

Analysis of Explanation 1

Table 3 presents the development of Reut's first explanation during the discussion. We start by listing the knowledge elements that Reut used in her explanations:

> *Reut1A:* (1) A vacuum tends to be filled; (2) A vacuum needs to be reduced.
> *Reut2A:* (3) Particles are things (4) Things in the vicinity of vacuum will eventually enter in to fill it.
> *Reut 3A:* (1) A vacuum tends to be filled

Intuition

We argue that the core of this explanation is an *e-prim* - "a vacuum has to be filled" - and that the knowledge elements # 1, 2, 4 are different manifestations of this. " A vacuum has to be filled" meets the first four mandatory criteria in Table 1 for identifying an e-prim: *Functionality*: it explains to Reut why the bottle shrank. *Obviousness*: Reut says: "I don't know why this happens, but I know it does… I even don't know why I know it" (Reut1b), "this is what I know" (Reut3c). *Developmental history:* source in language - the phrase "a vacuum will always be filled" is a common idiom in Hebrew (a nearly equivalent idiom in English is "nature abhors a vacuum"), source in interactions in the physical world - e.g., drinking from a straw, playing with a syringe. *Triangulation of expression:* KE 1, 2, 4 are different manifestations of the same e-prim. *Triangulation of form and content* - the notion that a vacuum has to be filled is mentioned in other studies. For instance Wiser and Smith (2008) discuss "a deeply held metaphysical principle that vacuum does not exist in nature" (p. 220), and diSessa (1993) identifies "vacuum impels" as a p-prim, but admits that this p-prim might be cultural to some extent. Given the fact that all five criteria were met we assumed that "a vacuum has to be filled" is an e-prim in Reut's knowledge system.

Local Coherence

All the knowledge elements that Reut activated (1, 2, 3, 4) locally cohere with one another and suggest coherent inferences that can explain the observed phenomenon in a convincing way: The bottle shrunk to fill the vacuum because the vacuum had to be filled (Reut1c).

Mechanistic Reasoning

To establish this measure we used the codes listed in Table 2 (Russ et al., 2008).

> *Reut1a:* (1) Describing the target phenomena ("the bottle shrunk"); (3) identifying entities: ("vacuum");
> *Reut1b:* (3) identifying entities ("vacuum", "vessel", "particles or something"); (4) identifying activities ("if there is a vessel that has vacuum inside it and there are particles outside, or something… it always enters the vacuum to fill it"); (5) identifying properties (the property of vacuum) (6) Identifying the organization of entities (the vacuum is inside the vessel the particles are outside the vessel).
> *Reut1c:* (1) Describing the target phenomenon ("the bottle shrunk"); (3) identifying entities: ("vacuum"); (5) identifying properties (the property of vacuum - "if there is a vacuum so something will fill it"), (6) Identifying organization of entities ("*here* there is more than...").

We argue that there is no chaining backwards (7) in this utterance, even though the structure of Reut's1c utterance approximates chaining backwards. We use the definition of chaining backwards in the following manner: using knowledge about the causal structure of the world (the property of a vacuum that it has to be filled) to make the claim (the bottle shrank to fill the vacuum) about what brings about the current state (the bottle has shrunk). In our view, this would be valid only if the bottle's "walls" were explicitly identified as an entity, but they were not, and hence Reut's argument is circular. Without this identification, chaining is not part of Reut's explanation.

Summary

All in all, we argue that it is not surprising that Reut did not immediately abandon the explanation she generated. It was based on a strong intuition, it presented local coherence that was developed during the three utterances, and its causal strength was not complete but was quite impressive, since apart from chaining, almost all the components of mechanistic reasoning seemed to be present (#1, 3, 4, 5, 6 in Table 2).

Explanation 2

The interviewer asked Reut: "Can you repeat what you understood *[from Natalie 1D]*? Table 4 presents the conversation between Natalie, Reut and the interviewer in which Reut constructed her second explanation, which started as a mere repetition of Natalie's idea and gradually developed and became "her own".

Table 4: Reut's second explanation

Reut2a:	What I understood... like, what I understood - there is the bottle… now… hm... one second *[looks to the side]*... first the particles move, right? So what happens is that they bounce into the bottle and they push it inside *[gestures pushing with her hand]*. Now, before, there were more, there were many particles…
Int.:	Can you explain it using one of the drawings? Try doing it here *[gives them a blank page]*
Reut2b:	*[Reut and Natalie take the paper]* Can I try on this? Say this is a bottle *[draws a rectangle]* lovely *[laughs]* now what is happening, we have plenty of particles here that push in this direction, and try.... to push it out *[gesturing "out" with her hand]* and particles from this direction that try to push it inside *[draw arrows on the page and gesture with her hand "inside"]*. This is what I understood, this is what I know.
Natalie:	And now with the pump, we took out, like,
Reut2c:	Yeah
Natalie:	Lots of particle from here
Reut2d:	So now we have here, like just two particles...
Natalie:	More... more force pushing from the outside
Reut2e:	And less force to resist, so it shrinks. It has nothing to do with a vacuum.

Reut Prefers the Second Explanation Over the First

Note that in the second part of utterance Reut2e, Reut states "It has nothing to do with a vacuum." Thus she rejected her previous explanation and accepted the current one.

Analysis of Explanation 2

Table 4 presents the development of Reut's first explanation during the discussion. Before we start, we would like to note the interviewer response when Reut said that she had no idea whether Natalie's explanation is correct or not. The interviewer asked her to restate it. During her restatement and elaboration of this explanation she changed her mind and became convinced of the explanation that *she* constructed during this process. Again, we start by listing the knowledge elements that Reut used in her explanations:

Reut2a: (1) the particles move; (2) the particles bounce on the walls of the bottle and push it; (3) particles outside of the bottle push the bottle's walls inside (4) pumping takes the particles out of the bottle.
Reut2b: (2) the particles bounce on the walls of the bottle and push it; (3) particles outside of the bottle push the bottle's walls inside; (5) particles in the bottle push the bottle's walls outside.
Reut2c: (4) pumping takes the particles out of the bottle.
Reut2d: (4) pumping takes the particles out of the bottle.
Reut2e: (6) fewer particles means a smaller pushing force; (7) the stronger force wins (overcoming p-prim; diSessa, 1993).

Intuition

At the core of Reut conviction lies the p-prim - "overcoming" - the intuitive schematization of one force or influence that overpowers another (diSessa, 1993). Note that Reut did not start the explanation with conviction. At first she was merely reconstructing Netalie's explanation, and her use of the word "push" (Reut2A) was to reference Natalie's explanation (Natalie1d). But in the next utterance Reut discussed conflicting forces: "we have plenty of particles here that push in this direction, and try.... to push it out *[gesturing "out" with her hand]*

and particles from this direction that try to push it inside *[draw arrows on the page and gesture with her hand "inside"]* (Reut2b). Note the triangulation of expressions for conflicting forces. Not only did Reut verbalize the conflicting forces, she also simulated them with her hands. P-prims are non verbally encoded, and it may well be that the gestures imitating the conflicting forces activated the notion of overcoming. After Reut explicitly expressed how pumping will result in fewer particles in (Reut2c & Reut2d, KE #4) and that fewer bouncing particles means less force (Reut2E, KE#6), the overcoming p-prim was activated, and was used with strong conviction.

Note that the reliability priority of overcoming is much higher than that of the e-prim "a vacuum has to be filled", and after it was activated Reut concluded that the shrinking of the bottle "has nothing to do with a vacuum" (Reut2e). The differences in the reliability priority could be attributed to the fact that e-prims that are also p-prims are abstracted from the very basic and frequent interactions with the physical world, whereas the roots of an e-prim such as "a vacuum has to be filled" are not as deep.

Coherence

The knowledge elements that Reut activated (#1-7) locally cohere and support one another, suggesting coherent inferences that can explain the observed phenomenon in a convincing way: The bottle shrunk since the particles out of the bottle pushed on the bottle harder than the opposing push of the particles in the bottle (Reut2e).

Mechanistic Reasoning

Reut2a: (3) Identifying entities (bottle, particles); (4) Identifying activities (the particles bounce, push); (5) Identifying properties of entities (particles move); (B) Animated model (Reut gestures with her hand to mimic the pushing of the particles)

Reut2b: (3) Identifying entities (bottle, particles); (4) Identifying activities (the particles push); (6) Identifying organization of entities (particles outside and inside of the bottle); (B) Animated model (Reut gestures with her hand to animate the conflicting directions of the particles being pushed in and out of the bottle)

Reut2c: (2) Identifying setup conditions (when we start pumping we take particles out of the bottle); (6) Identifying organization of entities (particles outside and inside)

Reut2d: (2) Identifying setup conditions (due to the pumping there are very few particles left in the bottle); (6) Identifying organization of entities ("we have *here* like two particles")

Reut2e: (7) Chaining backwards - using knowledge about the causal structure of the world (pumping the bottle took out particles from the bottle, thus fewer particles hit the inside of the bottle relative to those that hit the outer walls) to make claims (the force applied by more particles outside the bottle, overcame force applied by fewer particles inside the bottle) about what brought about the current state (thus, the bottle has shrunk).

Summary

Reut changed her mind and preferred the second explanation to the first one. This was accounted for in terms of the three dimensions of the internal metric for certainty in explanations. *Intuition:* The second explanation activated a stronger intuition than the first one (a p-prim vs. "just" an e-prim). *Local coherence:* The degree of coherence of the second explanation was higher than the first, since more knowledge elements cohered and supported one another. *Mechanistic reasoning:* The degree of mechanistic reasoning was higher since *all* the components of mechanistic reasoning were present.

Conclusion

In this paper we attempted to highlight students' internal evaluation of explanations, as opposed to external evaluations of explanations that are based on compatibility to scientific formulations. We explored the sources of students' sense of conviction of their own explanations and what guides their preferences for one explanation over another. We described the process of internal evaluation with a metaphorical theoretical construct *the internal metric for certainty in explanation*. The main contribution of this paper is the operationalization of this internal metric into empirical traceable terms. We envision the internal metric as composed of several dimensions. In the case of the internal evaluation of explanations of phenomena in the physical world, we suggested three dominant dimensions that form the internal metric: intuition, local coherence, and mechanistic reasoning. We operationalized each dimension and specified criteria for its evaluation based on students' self generated explanations. The illustrative analysis shows that the framework can explain conviction in an explanation, as well as relative conviction and the preference of one explanation over another.

It is important to note that paths of developing explanations guided by external evaluation and internal evaluation may not necessarily converge. This raises the the issue of the educational merits of the ability to externally measure students' internal convictions. We argue that students' agency in the developmental path from intuitive to scientific explanation is crucial for maintaining a sense of understanding, and hence the ability to assess students' convictions in explanations is a key component in instructional design.

References

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 421-441.

Berk, L. E. (1985). Why children talk to themselves. *Young Children, 40*(5), 46-52.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science, 13*(2), 145-182.

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

diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction, 10*(2&3), 105-225.

diSessa, A. A. (forthcoming). The Construction of Causal Schemes: Learning Mechanisms at the Knowledge Level. *Cognitive science*.

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

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

Gentner, D., & Gentner, D. R. (1983). Flowing waters or moving crowd: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 99-130). Hilisdale, NJ: Lawrence Eribaum.

Gopnik, A. (1998). Explanation as Orgasm*. *Minds and machines, 8*(1), 101-118.

Harman, G. H. (1965). The inference to the best explanation. *The Philosophical Review, 74*(1), 88-95.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps : analogy in creative thought*. Cambridge, MA: MIT Press.

Inagaki, K., & Hatano, G. (2002). *Young children's thinking about biological world*. New York, NY: Psychology Press.

Kapon, S., & diSessa, A. A. (2012). Reasoning Through Instructional Analogies. *Cognition and Instruction, 30*(3), 261-310.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. Hirschfield & S. Gelman (Eds.), *apping the mind: Domain specificity in cognition and culture* (pp. 234-254). Cambridge, UK: Cambridge University Press.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in cognitive sciences, 7*(8), 368-373.

Keil, F. C. (2006). Explanation and understanding. *Annual review of psychology, 57*, 227-254.

Keil, F. C., Rozenblit, L., & Mills, C. (2004). What lies beneath? Understanding the limits of understanding. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 227-249). Cambridge, Mass: MIT Press.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of science*, 1-25.

Metz, K. E. (1991). Development of explanation: Incremental and fundamental change in children's physics knowledge. *Journal of Research in Science Teaching, 28*(9), 785-797.

Mohan, L., Chen, J., & Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46*(6), 675-698.

Parnafes, O. (2012). Developing Explanations and Developing Understanding: Students Explain the Phases of the Moon Using Visual Representations. *Cognition and Instruction, 30*(4), 359-403.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science education, 92*(3), 499-525.

Salmon, W. C. (1989). Four decade of scientific explanation. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation, Minnesota studies in the philosophy of science* (Vol. 13, pp. 3-219). Minneapolis: University of Minnesota Press.

Toulmin, S. E. (1972). *Human understanding*. Princeton, NJ: Princeton University Press.

Wiser, M., & Smith, C. (2008). Learning and teaching about matter in grades K-8: When should the atomic-molecular theory be introduced? In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 205 – 237). NY: Routledge.

Acknowledgments

We would like to thank Rosemary Russ for a useful discussion about the application of the framework of mechanistic reasoning.