CHAPTER 6: RESULTS -- FACTORS INFLUENCING SENSE-MAKING DISCUSSION

Chapter 5 set the stage for this chapter by detailing the quantitative differences in sense-making discussions in cycles 3-5 in terms of distributions of sense-making instances and the percentages of time that groups dedicated to sense-making discussion. This chapter explains those differences in terms of personal, task, group, and contextual factors, thereby answering the final research question of this study:

Which factors provide support for students' sense-making discussion?

Below, Results 1-8 from chapter 5 are explained in numerical order. Result 1 is broken into two sub-results: 1a (dealing with clarifications) and 1b (dealing with descriptions/definitions/connections). But before I start explaining these eight results, one concern that needs to be addressed is whether my assumption about the expected distribution of sense-making instances was justified.

Testing an Assumption: Was Sense-Making Discussion Primarily Driven by the Written Curriculum?

A discussion about instances of SMD being "higher" or "lower" than expected only makes sense if the expected distribution is reasonable. In chapters 3 and 5, my assumption was that the expected distribution should be based on the distribution of sense-making instances in the written curriculum, because of the further assumption that the groups' SMD would primarily be curriculum-driven. In fact, this turned out to be the case. Groups rarely engaged in SMD outside of answering worksheet questions or solving worksheet problems.

Overall, 99.5% of verbal sense-making instances (523 out of 526) were directly or indirectly prompted by questions, directives (e.g. "Write your explanation here"), and graphing/diagramming activities in the curriculum materials. In other words, when the materials asked the students to engage in a thinking and writing activity, students typically engaged in conversations (mostly brief, some lengthy) about the required explanations, diagrams, etc.

The following discussion (group 1, from activity 2 in cycle 5, Exploring What Causes Gravity) has a discourse structure typical of the SMD in this study. Its structure is typical in that one student reads aloud from the materials, another student responds to the reading, and then further conversation ensues.

D (reading) "Do you think that the earth's magnetism …?" (remainder of the written question: "…causes gravity?") I don't think so.

L I don't either.

G Do we have magnets in our bodies? No.

D That's the thing. How does it hold things down if not all things are magnetized?

G We'd be floating around.

D Yeah. (laughs) Ok, so we don't think that magnetism causes gravity -- right? (D, L write answers in their notebooks)

L Because not everything can be magnetized.

In this example, the question under discussion was: "Do you think that the earth's magnetism causes gravity? What is your evidence?" (Only the first part of the question was read aloud.) Immediately preceding the discussion, the group determined the magnetic and gravitational properties of various materials. The purpose of the activity was to demonstrate that not all objects are magnetic, yet all objects are affected by gravity.

In this example, Darla read aloud a portion of the question, and then answered the question verbally. Then three (of four) group members engaged in SMD. In discourse terms, Darla's reading of the question served as an initiation of discussion, as did her spoken answer to the question. Other statements were meant to extend or qualify previous contributions to the conversation. The main point of the example, however, is that the discussion owed its presence to a question in the materials. This sense-making exchange is extremely typical in this regard.

Of course, the overall distribution of sense-making instances in group discussions is expected to deviate from the distribution of sense-making components in the written curriculum. After all, groups do not talk about each question. Also, questions requiring a particular component of sense-making do not lead to discussions composed solely of that component. For example, prediction questions do not lead to discussions composed solely of predictions. Ideally, a prediction from one group member (we hope) elicits clarifications, requests for evidence, and statements of basic principles of physics from the other group members. The same could be said about "explanation" questions, "clarifying" questions, and so forth. Still, the fact remains that radical differences between the actual and expected distributions are in need of explanation. Such differences existed between two of the components of verbal sense-making: clarifying the facts of a phenomenon or result, and defining, describing, and connecting scientific concepts, procedures, processes, and representations.

Result 1a: Clarifications

<u>Result 1a: There were more instances of clarifying the facts of a phenomenon or experimental result than expected.</u>

Relevant explanations and assertions:

- The number of clarifying instances was higher than expected because group 2 had difficulty coming to agreement on certain basic facts about gravity, instantaneous speed, friction, and air drag.
- The unexpected clarifying discussions were necessary because they explored conceptual details that were not specifically addressed by the CIPS curriculum.

Looking at Appendix 4, certain activity sub-sections were noticeably above average in terms of numbers of clarifications: activity 3 (Slowing Down) and activity 5 (What's a Little Drag?) in cycle 4, and activity 2 (Exploring What Causes Gravity) in cycle 5

Careful analysis of the transcripts for these sub-sections shows that the clarifications were due to group 2 unexpectedly going off on clarifying tangents about air drag, gravity, friction, and instantaneous speed, where "unexpectedly" means that the clarifications were aimed at levels or aspects of physics understanding that were beyond the scope of the CIPS activities. The drag, gravity, and speed issues that came up (listed below) were thought to be too specific or too detailed to be explicitly addressed by a middle school physical science curriculum, hence the "unexpected" nature of the student clarifications on these topics. Regardless, it was these unexpected discussions concerning the details of drag, friction, gravity, and instantaneous speed that skewed the distribution of instances in favor of "clarifications of facts" in Table 5-1.

Specifically, the unexpected clarifying discussions were focused on answering the following questions, respectively, about force and motion:

- Does a slowing object (due to friction) move at constant speed and then slow down -- or does it just slow down?
- Are drag and air friction the same thing?
- Is there gravity in earth's atmosphere and in space?

Friction and instantaneous speed. In activity 3 of cycle 4 (Slowing Down), groups were asked to: shove a wood block, release it, watch the block's motion after the shove, and then describe the block's motion after the shove as either speeding up, slowing down, or moving at constant speed.

Group 2 chose to stray from the suggested "speeding up, slowing down, or moving at constant speed" descriptions of motion, however, because some group members (Roxanne, Sabrina) felt that another option might be correct: an initial movement of constant speed, and then a slowing down motion. This idea of "constant speed, then slowing down" led to almost 4 minutes of heated, in-depth discussion between the group members, in which they tried to clarify whether there was an instantaneous slowing at every point (my words, not theirs) or whether there was a portion of the block's motion where the speed was unchanging. (During this discussion, there was a good deal of vagueness about the term "constant", as the group members used the phrase "constant speed -- for maybe a millisecond" to refer to the peak instantaneous speed on the speed graph that Arthur had drawn to clarify things.) In fact, group 2 spent so much time discussing these clarifying issues that they weren't able to do any of the required energy and speed diagrams in the activity. One likely explanation for the conceptual confusion in this case is that the previous activities only addressed four types of motion: constant speed, speeding up, slowing down, and zero speed. The idea of instantaneous velocity was never mentioned in class or in the curriculum materials, and so students were left to themselves to grapple with the issue when it came up (as it did here).

Air drag and friction. Most science students have difficulty understanding the differences between drag and friction ("air friction", especially), including those students in high school and college. In CIPS, drag and friction are presented in the following manner: drag and friction are both resistive interactions, meaning that they can slow objects down if acting alone; drag comes from an interaction between air and a "dragging" object (a sail, drag parachute, etc.), and friction comes from an object rubbing against a surface. On the surface, these concepts are extremely similar. Both friction and air drag slow things down, and both involve pushing/rubbing (an unclear distinction, at best). At the middle school level, the difficulty in adequately presenting the differences between friction and drag is that molecule- or atomic-level explanations are needed to properly explain the differences between the two[1] -- and students have not yet been exposed to the atomic/molecular nature of matter.

The general (i.e., non-atomic) differentiation in CIPS between air drag and friction is precisely the issue that led to group 2's unexpected clarifying discussion in activity 5 of cycle 4 (What's a Little Drag?). During the discussion, in which students debated the motion of two carts (one with a sail, one without), Roxanne simply did not recognize that there was a difference between the two phenomena (she would use the word "friction" to describe the air's interaction with the cart's sail), while Arthur repeatedly tried to get the

[1] The difference is that energy (heat) transfer from friction increases the vibrational energy of the rubbed surfaces, while the energy transfer in drag phenomena results in an increase in the air molecules' translational energy (i.e., the air molecules acquire an increased translational speed).

point across that friction was not a concern, and that drag was the relevant phenomenon in this instance.

Gravity. The last conceptual issue that resulted in an unexpected clarifying discussion was the following: Does gravity exist away from the surface of the earth (in the atmosphere, in space)? In activity 2 of cycle 5 (Exploring What Causes Gravity), for example -- an activity that was only supposed to focus on the effects of the earth's magnetic field, spin, and atmosphere on earth's gravity -- Sabrina argued that there was no gravity in the earth's atmosphere or in space (the sun had no gravity, she argued), while Arthur tried to convince her otherwise. One of the issues here is that students were not really given the opportunity to come to agreement at the classroom level (in a whole-class discussion, for example) on the precise "'location" of gravity before diving into its possible cause. Students understood that there was gravity near the surface of the earth, but many had problems understanding that gravitational forces exist a significant distance away from the earth's surface (in the earth's atmosphere, e.g.) -- including outer space.

Typically, CIPS activities gave students the opportunity to agree on the facts of a phenomenon before trying to explain that phenomenon. Gravity was one of the few exceptions in this regard, which was intentional on the part of the developers. Ideally, students were to do the following, in order: a) understand the cause of gravity on earth, and b) understand the cause of gravity in general. This was (and is) a good idea, since students were given the chance to understand an everyday phenomenon (earth's gravity) and then abstract and generalize that understanding to define a general principle (the universal law of gravity). The only apparent drawback of this ordering is that, as seen here, some students may wish to address the general case of gravity earlier than the curriculum allows, as they did in this study. It was this clarifying discussion about the universal nature of gravity that made the last significant contribution to the unexpected number of clarifications in the distribution of sense-making instances.

Result 1b: Descriptions, Definitions, and Connections

Result 1b: There were fewer instances of describing/defining/connecting scientific concepts, procedures, processes, and representations than expected.

Relevant explanations and assertions:

- The number of describing/defining/connecting instances was lower than expected because: 1) "Making sense" and "Now what does your team think?" sub-sections were frequently skipped by at least one group, 2) group 2 students worked on one "Making sense" sub-section individually, instead of as a group, and 3) "Prepare your wipe board!" sub-sections contained less sense-making than expected.

Each of these relevant factors is described in detail below.

Reason 1 for DDC being less common than expected: "Making sense" and "Now what does your team think?" sub-sections were frequently skipped by at least one group. A basic assumption behind the use of the written curriculum as an approximation for the distribution of verbal sense-making instances was that both groups would spend time in each activity sub-section. The "expected" percentages in Tables 3-6 and 5-4 reflect this assumption. To my chagrin, one or both groups failed to spend time in 13 of the 45 sub-sections in cycles 3-5 of the Force and Motion unit. Two of the eight "Making sense" sub-sections were skipped by one group, and two were skipped by both groups. One of the three "Now what does your team think?" sub-sections was skipped by one group, and two were skipped by both groups.

The result of skipping these sub-sections is that groups did not have the opportunity to engage in the sense-making supported in the sections. The purpose of the "Making sense" sub-sections was for students to reflect back on the results of an experiment to determine whether the results supported or refuted a particular scientific hypothesis (e.g., magnetism causes gravity)-- a process that involves a good deal of defining, describing, and connecting among scientific concepts, procedures, processes, and representations. The purpose of the "Now what does your team think?" sub-sections was for students to rethink the pre-consensus status of their explanatory ideas -- a process which, again, involves defining/describing/connecting. Taken together, the net result of these "skippings" is that students didn't engage as frequently as expected in the processes of reflecting and rethinking, and so the amount of defining/describing/connecting turned out to be far less than what was hoped for (illustrated in Table 5-4).

Why were these sub-sections skipped? Three reasons: lack of time, lack of student interest or awareness, and lack of formal teacher support.

When the overall structure of CIPS was formulated, common sense dictated that the necessary (but unfortunate, it turns out) placement of a "Making sense" sub-section would be at the end of each Development activity. In these activities, groups were to first engage in hands-on science ("What really happens?") and then reflect on the conceptual relevance of the experimental results ("Making sense"). What developers had not foreseen was that most experiments would go over their developer-allotted time limit, and so the final "Making sense" sub-sections in these activities suffered by either being skipped entirely or by being dedicated only a few minutes of class time. This occurred because CIPS, at the time of this study, was in its initial pilot year; developers had set careful time guidelines for each activity and activity sub-section, but -- almost universally -- classes spent more time than expected on each sub-section.[2] Due to the developers' inexperience in these matters (myself included), the suggested time for each sub-section was simply unrealistic. Time turned out to be such a concern, in fact, that -- in the interest

[2] At the time of this writing, CIPS is in its second year -- and time is still a huge problem. Activities are still taking longer than expected, which means that most activities will have to be shortened, and some activities will have to be deleted entirely.

of finishing all 5 curriculum units by the end of the year -- the teacher would occasionally start the next day with a new activity even when the "Making sense" sub-section from the previous activity had been skipped.

One example of skipping a "Making sense" sub-section during a transition between activities was in the transition between activity 5 (What's a Little Drag?) and activity 6 (Putting it All Together) of cycle 4. The class only made it as far as the simulations portion of activity 5 (the simulations come just before "Making sense"), and the teacher started with activity 6 the next day without going back to complete "Making sense" in activity 5.

The other sub-section that was effectively skipped for time reasons was the "Making sense" section of activity 2 in cycle 5 (Exploring What Causes Gravity). With 3 minutes remaining, students in class 2 were instructed to finish the remainder of the activity: the last question in the gravity/air pressure experiment, the "Making sense" section, and the "Idea journal" section. Because group 2 only had time to answer the last question in the gravity/air pressure experiment, the group never made it to the "Making sense" section.

It's one thing for a group to justifiably skip a section because the entire class spent too much time on earlier portions of the activity. It's another for a group to skip a section because they wasted so much time goofing off and waiting around for unnecessary instructions that they couldn't finish the last few activity sub-sections, even though most other groups were able to finish those sections. Unfortunately, group 2 skipped "Making sense" in activity 3 of cycle 3 (Lots of Pushes and Pulls!) precisely for these reasons. The group demonstrated its general disinterest in activity 3 by being off-task for a good portion in the activity (putting it way behind in comparison with other groups in the class), and then the group only wasted more time when, after finishing p. 3 of the activity, the group waited for the teacher to tell them to move on to p. 4 -- even though the teacher had already made numerous announcements to the class that they should finish the entire activity. The overall result of the group's lack of interest and awareness was that Sabrina, Arthur, and Roxanne barely had time to finish the graphing activity on p. 4, let alone get to (and complete) the "Making sense" section on p. 5.

The last reason why groups sometimes skipped the sub-sections aimed at eliciting defining/describing/connecting discussions sections ("Making sense", "Now what does your team think?") was a consistent lack of formal support. This lack of support was consistent for "Now what does your team think?" sub-sections, but it only occurred once for "Making sense". The one time that the teacher did not formally support "Making sense" was in activity 3 in cycle 4 (Slowing Down), where both groups skipped the sub-section because, after groups completed the Sliding Blocks experiment, the teacher directed the groups to finish the "Idea journal" section without also telling groups to complete "Making sense". It should therefore not come as a surprise that both groups skipped the "Making sense" sub-section in this activity. (Why the teacher failed to direct the students to complete "Making sense" is unclear. He may have simply forgotten, or he may have been concerned about time.) The most consistent lack of teacher support was for the "Now what does your team think?" sub-sections. Ideally, in each consensus activity, groups were to formally revisit their ideas in "Now what does your team think?" -- at which time they could make changes, etc. -- and then they were to record those ideas in their final form during "Prepare your wipe board!" In practice, group 2 never spent time on a "Now what does your team think?" section, and the one time that group 1 spent time on the section (in activity 5, cycle 5) was purely by accident.[3]

The way that the teacher set up each consensus activity was to say something like the following:

"Get your wipe boards. You have 5 minutes to get your Idea Log presentations ready." (Putting it All Together, cycle 3; class 1)

"Review the three ideas, and decide what you're going to write…I want the presentation boards up in ten minutes." (Putting It All Together, cycle 4; class 2)

"Summarize evidence and ideas for ideas 1 and 2…Let's get to work and see if we can get the presentations up and running." (Putting it All Together, cycle 5; class 1)

The teacher's directions for the cycle 3 and cycle 5 consensus activities were clearly focused on putting up the presentation boards as quickly as

[3] Atypically, the group had forgotten to record an idea in their Idea Journal. They then spent

possible. While the cycle 4 instructions superficially directed the groups to spend time on the "Now what does your team think?" sub-section (the teacher's words: "Review the three ideas…"), what actually happened was that, as soon as the teacher finished his instructions, students leapt up to get the presentation boards -- and the teacher did nothing to stop them. In actions or in words, the teacher's instructions to the class for the cycle 3-5 consensus activities did not formally support the revisiting of the cycle ideas that was supposed to occur in "Now what does your team think?", and so these sections were frequently skipped. The result of this was that groups did not spend as much time on defining/describing/connecting as expected.

Reason 2 for DDC being less common than expected: Group 2 worked on one "Making sense" sub-section individually, instead of as a group. In the spirit of both individual and group sense-making, CIPS activities were designed such that certain sub-sections were always to be done individually (e.g., "What do you think?" and "Idea journal" sub-sections) and certain sub-sections were always to be done in groups (e.g., "Making sense" and "What really happens?" sub-sections). Any deviation from this consistent structuring of sub-section types would therefore be a surprising result to the curriculum developers, who had no reason to think that this individual and group structuring of sub-sections would be compromised. Yet, the record of class activity in the Force and Motion unit (videotapes, fieldnotes etc.) clearly shows

---

time on "Now what ideas does your team have?" to copy the missing idea from earlier work.

that sub-sections that were to be done individually were occasionally done in groups, and vice versa. All told, there was one instance where an "individual"-type sub-section was done in groups, and three instances where "group"-type sub-sections were done individually. One of the latter cases happened to occur in a "Making sense" sub-section; it was this case that negatively influenced the smaller-than-expected number of describing/describing/connecting instances in the overall distribution.

The first case of group/individual confusion occurred in the "What do you think?" sub-section of activity 3 in cycle 3 (Lots of Pushes and Pulls!). "What do you think?" sub-sections were meant to be places where students make individual predictions about a phenomenon, but the teacher explicitly instructed the students (in both classes) that this section of activity 3 should be filled out by "you and your team". Another instance occurred later in that same activity, in "Let's find out!", where students were to draw force diagrams in groups; instead, the teacher's parting "Good luck" at the end of his instructions (in both classes) implied that the diagramming was a test-like activity that should be done individually and silently. (At one point, Sabrina asked a fellow group member: "Do we do this on our own?") The third instance occurred when, in activity 2 of cycle 5 (Exploring What Causes Gravity), the teacher told class 1 that they needed to fill out their predictions about the effect of air pressure on gravity before they could watch the air pressure video, and so students dutifully (silently, individually) recorded their predictions. (This was supposed to have been a group activity, but the teacher did not encourage discussion.) Finally, in the "Making sense" sub-section of activity 3 of cycle 3 (Lots of Pushes and Pulls!) -- the instance of confusion that is relevant to there being fewer descriptions, definitions, connections than expected -- group 2 was unclear as to whether "Making sense" in this activity should be done individually or as a group. Roxanne and Sabrina debated the matter, and eventually came to the conclusion that the section was to be done individually. The result was that group 2 worked silently for the first three minutes of the section, although they eventually backslid (or so they thought) and discussed their ideas for approximately 30 seconds as a group.

Reason 3 for DDC being less common than expected: "Prepare your wipe board!" sub-sections contained less sense-making than expected.

The factors explaining this result are found in Result 3, below. To briefly summarize, groups didn't spend much time on sense-making in these sub-sections because the revisiting and reconstruction of ideas over the course of a curriculum cycle did not occur as planned. Instead, students typically were satisfied with a particular idea long before the consensus activity arrived, and so the students did not feel the need to revisit, and therefore re-discuss, these ideas in the consensus "Prepare your wipe board!" sub-sections. Most of the time in these sections was dedicated to students dictating their already-finalized ideas to the person doing the writing on the presentation board.

Result 2: Predictions and Evidence

Result 2: More than half of the predictions were not supported by evidence.

Relevant explanations and assertions:

- Predictions (both level 1 and level 2) were the direct result of "what do you predict?" questions in the curriculum. Level 2 predictions included supportive reasoning; level 1 predictions did not.
- Four prediction questions did not ask students to record or discuss their supportive reasoning. These questions elicited level 1 predictions (i.e., predictions without evidence). This may not be a causal relationship, however, due to a confounding effect.
- Overall, there were a number of reasons why a student might not discuss his or her reasoning behind a prediction: 1) the prediction was too simple, and so the student's reasoning was obvious enough that it didn't need to be stated, 2) the student misinterpreted the nature of the prediction, 3) the student was interrupted, or 4) the student's prediction was a guess (i.e., there was no clear reasoning behind the prediction).

On the surface, the existence of more level 1 predictions than level 2 predictions appears to reflect the students' failure to draw on supportive reasoning for their predictions. The concern would then be that the students, because of their possible reliance on guesswork, would not be in a position to compare experimental results to their predictions and consequently determine if their original reasoning and ideas are still valid, or possibly in need of change. However, in practice, most level 1 predictions in cycles 3-5 were not the result of guesswork. It appears that, in general, students did draw on supportive reasoning for their level 1 predictions -- they just didn't bother to explicitly communicate that reasoning to their teammates. This occurred for a variety of reasons, all of which are discussed below.

Before addressing the various reasons for the relatively high number of level 1 predictions, let me first repeat one of the general sense-making results in this study (from Result 1, above). Overwhelmingly, students only engaged in sense-making discussion when they ran across questions, directives, or graphing/diagramming activities that required sense-making. This result holds for predictions as well. Students were not seen to make spontaneous predictions during experiments, for example. They only made verbal predictions when the materials required students to record or discuss their predictions. Recognizing this effect, a logical hypothesis to pursue is whether students only provide evidence for their predictions when asked to do so.

Lack of supportive reasoning: Curriculum prompts.

In this study, it is difficult to address the issue of whether students only discuss their supportive reasoning when prompted to record/discuss their reasoning. There are two reasons why this is so. First, there were only four instances in the curriculum where students were not asked to provide reasoning for their written predictions. This gives us an extremely small "negative" sample to work with. Second, the prediction questions that failed to prompt students for their supportive reasoning happened to focus on very simple predictions. In these instances, the reasoning was obvious enough that it didn't need to be communicated to the rest of the group. (This is an effect that is discussed in more detail below.)

Therefore, even though the four promptless prediction questions failed to elicit anything but level 1 predictions, no firm conclusions can be drawn about the relationship between the lack of "evidence" discussion and the lack of curricular prompting for evidence. On the other hand, there are a number of factors affecting the voicing of supportive reasoning that can still be investigated.

Lack of supportive reasoning: Predictions that are too simple.

In activity 1 of cycle 3 (Can You Lend Me a Hand?), groups were asked to weigh two separate objects and then predict the combined weight of both objects. Group 1 (Lacey, Darla, Grace, Porter) found that the objects weighed two and three newtons, respectively, and then had the following brief discussion.

L Ok, (reads from worksheet) "predict what will the scale read when both objects...?" (confidently) It'll be 5 newtons.

D Five. Everybody agree with that?

L Do it! Ok, let's do it you guys! Let's do it! (weigh the objects)

Clearly, Lacey underwent the mental process of adding two and three newtons together -- that is, she assumed that the combination of pulls would be additive -- in order to predict what the combined weight would be. And since the basis for her prediction was fairly obvious, there was no real need to communicate that reasoning to her groupmates.

Group 2's level 1 predictions in cycle 3 (two of them) were also from activity 1, and were nearly identical to the example above. In both instances, students from the group implicitly used the idea that "combining is an additive process" to predict what the combined weight would be.

Another example comes from activity 2 in cycle 3 (Lots of Pulls!), in which students were supposed to "Predict what will happen and write your reasoning" before simulating a tug-of-war with lego bricks attached to a pulley. In the case where the number of blocks per side was equal, Lacey's verbal prediction was that "This is going to be even, I bet" -- with no mention of supportive reasoning. This was because the supportive reasoning was too obvious: the system won't move because there are the same number of blocks on each side. Common sense tells us that this will be the case.

The final example comes from activity 1 in cycle 4 (Will It Slow Down?), in which students were asked to consider whether a pushed skateboard (on earth), bicycle (on earth), shopping cart (on earth), and wrench (in space) would keep moving at a constant speed or eventually slow down. Universally, groups wrote on their presentation boards that the skateboard, bicycle, and shopping cart would eventually slow down, and the wrench would keep moving at constant speed. In groups 1 and 2 of this study, the resulting "prediction" conversations lasted less than 90 seconds, with the conversations primarily composed of an equal number of "slow down" (type 1) and "slow down because of friction" (type 2) predictions.

From a curricular standpoint, the problem with some of the "what does your team predict?" questions in the Force and Motion unit is that the correct predictions were arrived at too easily, primarily because of the simple numerical or "common sense" nature of the predictions. As soon as numbers were introduced into the prediction process (in the form of weights) in the cycle 3 activity, students naturally gravitated towards the basic mathematical operations (addition, subtraction, etc.) as a means of generating prediction, and then chose the only operation that could possibly make sense: addition. In short, the task was too simple and too procedural, and therefore individuals had no reason to announce their reasoning to the group. In the cycle 4 activity, the main "prediction" was to determine if objects slow down on earth -- a prediction that (in hindsight) nearly anyone of middle school age could make with great certainty.

Johnson and Johnson (reference) address these issues by stating that task goals for groups need to be appropriate for collaborative work. Put another way, the goals of group activities need to be complex enough so that group members are forced to rely on each other's ideas, knowledge, and experience in order to achieve the task goal. In the prediction processes outlined above, students could easily make the prediction on their own, and so there was no need for the students to engage each other in in-depth discussion.

Lack of supportive reasoning: Misinterpreting the nature of the prediction.

A fascinating phenomenon occurred in activity 2 of cycle 5 (Exploring What Causes Gravity). Twice, members of group 2 re-interpreted a hypothetical "if this were true, what would happen?" prediction as a simple "describe what will happen" prediction. The result of the re-interpretation is that the predictions were made too easy, and so students were less likely to vocalize their supportive reasoning.

The first instance of this phenomenon occurred when students were asked to test the possibility of the earth's spin causing gravity. To do this, they attached a lego to a pencil (with string), and then spun the pencil. The idea was for students to notice that the lego moves away from the pencil -- not toward it. (The analogy is that earth's gravity pushes us away from the surface of the earth, not towards the surface.) Before performing the experiment, groups had to answer a prediction question; the following is what transpired in group 2 at this point:

S (reads) "If gravity is caused by spinning, what should happen to the washer if you spin the pencil?"

R (spins lego) It should float, like this. (the lego goes horizontal)

In this example, the key result is that Roxanne did not appear to engage in the process of first making an assumption (gravity is caused by spinning) and then following that assumption to its logical conclusion (since gravity is caused by spinning, the lego should move toward the pencil). Instead, Roxanne seemed to skip the "assumption" aspect entirely and jump immediately into a prediction based on her past experience (the lego will move outwards, just as it always does.). Doing so caused Roxanne to avoid the real purpose of the prediction, which was to test an assumption about the cause of gravity. Instead, the group addressed a very different (and much simpler) purpose: to predict what happens when you spin a lego on a string.

Another example occurred 8 minutes later when students were asked to predict what would happen to an object's weight if it was put in a jar, and then the air was evacuated from the jar. (Afterwards, the class was shown a videotaped version of this demonstration. The result is that nothing happens to the object's weight; therefore, earth's atmosphere does not cause gravity.) The exact wording of the prediction question: "If gravity is caused by air pressing down on an object, what should happen to the object if the air is removed?"

S It will fall! I mean, it will float! It will float! Because, see...

A No it won't.

S ...there's air pressure pushing down on this, which is gravity. If there's nothing there, it will go whee-oo-eee-oo. (makes floating away motion with hands)

A Actually, no.

In this case, Sabrina tackled the prediction in the appropriate manner: she made the suggested assumption about gravity (atmosphere causes gravity), and then followed the assumption to its logical conclusion (if atmosphere causes gravity, the object will float when the air is removed). Arthur's response ("No it won't", a type 1 prediction) is not appropriate, because he has re-interpreted the prediction (as Roxanne did earlier) to simply predict what will happen -- a prediction that does not require supportive reasoning because of its intuitive and obvious nature.

An interesting question is why these re-interpretations occurred. On one hand, group 2 is notorious for being less-than-careful in reading and following written instructions. It is therefore possible that the group simply misread the questions. On the other hand, another explanation is that the "if this were true…" questions posed genuine cognitive difficulties for the students because of the fact that the questions forced students to follow complicated, abstract chains of logic. The hypothetico-deductive process of pretending that something is true and then making a prediction based on that imagined assumption is a process that many students find quite challenging

(Lawson, 1985). Perhaps the students could have used a bit more support from the teacher in terms of reading, understanding, and successfully answering this style of hypothetical question.

Lack of supportive reasoning: Interruptions.

Some predictions were labeled as type 1 predictions because the predicting students were cut off before they could finish their thoughts. The students may or may not have provided reasoning for the predictions, but interruptions prevented me from knowing one way or the other. As one might imagine, this reason for the existence of a type 1 prediction -- aside from concerns about student politeness and cooperation -- is slightly less interesting than the others.

Lack of supportive reasoning: Guesswork.

Although the general concern that students might have engaged in wholesale guesswork proved to be unfounded, I did come across a handful of verbal predictions that appeared to be guesses.

One such example is Grace's (group 2) response to a question in activity 3 of cycle 5 (More Exploration of What Causes Gravity) that asked students to consider, if hammers are dropped simultaneously on the earth and the moon (at the same height), which hammer might hit the ground first. The relevant discussion:

D (reads) "Which hammer will hit the ground first? Why? Write down each team member's name, idea, and reasoning Then, write the group's final idea after you have had a group discussion."

D (to G) Ok, Grace -- what do you think?

G What?

D It will drop. Which one will drop first?

G Teract's. (the hammer on the earth)

D (to G) Why?

G Because she's more closer to the...

D Because why?

G She's closer (to the ground). I don't know.

D Because she's closer to the ground? (G nods)

Grace's facial expression and tone -- coupled with her "I don't know" response to Darla's prompting for supportive evidence -- made it pretty clear that her initial prediction (the hammer on the earth would hit first) was actually a guess, or at the very least a prediction that had only been given a second or two of thought. As additional proof, Grace changed her answer less than 30 seconds later to suggest that the hammer on the moon might float, and therefore might not hit the ground at all.

Other clear examples of guesswork in CIPS are when Sabrina postulated that gravity was due to a gathering of friction at the surface of the earth (her comment: "I don't know if it's right. I just guessed."), and when --

after viewing a video on the Cavendish experiment --Arthur guessed that friction was the reason that two pencils on a table didn''t move closer together due to their mutual gravitational attraction (his comment: "But I really have no idea. That's just the first answer that came to mind.").

So, accepting that there were a few more examples like Grace's "hammer" prediction, it cannot be denied that students didn't always put forth the sort of rigorous cognitive effort that is necessary to make a scientific prediction based on sound supportive reasoning.

A possible confounding effect about the general nature of CIPS predictions is that students with guess-like predictions may have been less likely to vocalize their predictions in the first place -- and that therefore guess-like (nonverbal) predictions might have been more common than my verbal data would lead you to believe. In answering prediction questions in cycles 3-5, it is certainly true that one or two group members typically made level 2 predictions (predictions with evidence). But what about the silent members of the group? They may have been in unspoken agreement with the offered prediction and/or supportive reasoning, or it may simply have been the case that they had no idea what was going to happen -- which means that any prediction they <u>would</u> have made would have been a guess.

In the Force and Motion unit, a few prediction questions required each and every student to record their prediction/reasoning and/or share their prediction/reasoning with the group; based on what I saw on the videotapes, either requirement (recording, sharing) went a long way toward guaranteeing that each student went through the process of making an initial prediction based on reasonable evidence. However, as seen in the above "hammer" example with Grace, there were some instances when even explicit prompts for the sharing of ideas/reasoning couldn't prevent guesswork from some group members.

Result 3: Sense-making Differences between Sub-section Types

<u>Result 3: The percentage of time that groups dedicated to sense-making discussion was highest in "What does your team think?" and "Making sense" sub-section types and lowest in the "Prepare your wipe board!" sub-section type</u>.

Relevant explanations and assertions:

- "What does your team think?" and "Making sense" sub-sections were intended to elicit the most group sense-making discussion, so the relatively high percentages in these sub-section types are unsurprising.
- It was originally intended that CIPS students would engage in an extended metacognitive process of re-examining and re-developing their ideas over the course of each cycle. The culmination of this process was to occur in the consensus activity, at which point students were to engage in group and whole-class discussions in order to identify the ideas that best explain the cycle phenomena.

- This extended metacognitive process did not occur. Instead, students typically were satisfied with a particular idea long before the consensus activity arrived. There were three reasons for this: 1) group 1 thought that their ideas were basically correct, and no one in the group was in a position to provide the guidance necessary to clarify or improve these ideas, 2) group 2 didn't seem to value the ongoing self-examination of their ideas, and 3) the teacher would sometimes hint strongly at (and sometimes even explicitly tell) students the "correct" cycle idea.
- The result of students being satisfied with their ideas early in the sense-making process is that students did not feel the need to revisit, and therefore re-discuss, these ideas in the "Prepare your wipe board!" sub-sections in the consensus activities. Therefore, most of the time in these sections was dedicated to students dictating their already-finalized ideas to the person doing the writing on the presentation board. It was this typical behavior that played a major role in the low sense-making percentage in "Prepare your wipe board!" sub-sections.

Expectations: Sub-sections supporting significant sense-making. The "What does your team think?" and "Making sense" sub-section types were the sub-sections that were intended to elicit significant amounts of sense-making discussion, so the fact that these two sub-sections had the largest sense-making percentages indicates that the CIPS developers were fairly successful in achieving their goal.

"What does your team think?" sub-sections prompted SMD by having groups reflect on their initial ideas about an interesting phenomenon; they did so by having groups provide an underlying explanation for the phenomenon or make a prediction about the phenomenon. The "interesting phenomena" in these activities were a tug-of-war (cycle 3), friction (cycle 4), and gravity (cycle 5). The "Making sense" sub-sections, which always followed the "What really happens?" (i.e., hands-on experiment) sub-sections in the Development activities, supported sense-making discussion by prompting students to reflect on experimental results in order to rethink and reconceptualize the group ideas about pushes and pulls (cycle 3), resistive interactions (cycle 4), or gravity (cycle 5).

Since the "What does your team think?" and the "Making sense" results were already expected, perhaps the more interesting result is the comparatively poor faring of the "Prepare your wipe board!" sub-sections: on average, groups spent only 6% of their time on SMD in these sections.

Before examining this result, however, let us first recall the overall structure of the CIPS curriculum cycle and how the "Prepare your wipe board!" sub-sections were meant to fit into this overall structure.

<u>Expectations: "Prepare your wipe board!"</u>. As originally conceived, CIPS curriculum cycles were designed to support students' metacognitive

examination and restructuring of their physics ideas over the course of 5-7 days. The plan was for students to examine their initial ideas about a phenomenon (the Elicitation activity, day 1), after which the students were to revisit and modify their initial ideas by conducting experiments and reflecting on the experimental results (Development activities, lasting at least 3 days). Finally, students were to have small-group and whole-class discussions to decide which ideas best explain the phenomena in the cycle (the Consensus activity, the final day).[4]

"Prepare your wipe board!" sub-sections existed in two of these cycle activities: elicitation activities and consensus activities. In the elicitation activities, "Prepare your wipe board!" sub-sections were places where the groups were to finalize their initial predictions/explanations of a new phenomenon. They were to do so as they recorded these explanations/predictions, with evidence, on presentation boards. Similarly, in the consensus activities, "Prepare your wipe board!" sub-sections were places where groups were to come to final agreement on the group ideas, with evidence, that best explain the cycle phenomena (again, in the act of writing the ideas on presentation boards).

I emphasize that groups were intended to come to "final" agreement in "Prepare!" sub-sections because each elicitation and consensus activity already had a section that was dedicated to group discussion of the

[4] There were also application activities at the end of each cycle, but these activities were not

explanation/prediction or "best" ideas, respectively. In elicitation activities, the group discussion section was "What does your team think?"; in consensus activities, the group discussion section was "Now what does your team think?". These sub-sections immediately preceded the "Prepare your wipe board!" sections in elicitation and consensus activities, respectively -- and so, having already debated the ideas in a separate sub-section, groups should only have needed to finalize their ideas in the "Prepare your wipe board!" sub-sections.

<u>Explaining the result: Groups spent only 6% of their time engaged in sense-making discussion in "Prepare your wipe board!" sub-sections</u>.

A low percentage (even as low as 6%) is perhaps appropriate for sub-sections in which students had only been expected to put the finishing touches on their ideas. After all, much of the time in these sub-sections was necessarily spent on doing the actual writing on the presentation boards. There is one problem, though. As discussed above, "Now what does your team think?" sub-sections (in the consensus activities) were typically skipped by both groups. Accordingly, one would have expected the bulk of SMD in these activities to shift from the skipped "Now what does your team think?" sub-sections to the "Prepare your wipe board!" sub-sections. Did this occur? No. If we separate the "Prepare!" percentages into elicitation and consensus sub-sections, we find that the average for each is still 6%.

---

relevant to this study.

Given our increased expectation for SMD in the "Prepare!" sub-sections from the consensus activities -- due to the fact that students skipped the opportunity for SMD in the "Now what does your team think?" sub-sections -- why were the percentages in these sub-sections still a low 6%?

The main problem was that the original plan of week-long metacognitive restructuring didn't quite work as intended. Where the plan ultimately broke down was that, in practice, students did not wind up spending a significant amount of time revisiting and modifying their ideas. Why? Because groups were typically satisfied with their ideas long before the consensus activities rolled around. The conscientious group (group 1) didn't spend a lot of time revisiting their ideas because they thought that their ideas were basically correct, and no one in the group was in a position to provide the guidance necessary to clarify or improve these ideas. The not-so-conscientious group (group 2) didn't want to revisit their ideas because they didn't see any value in doing so. Also, occasionally, neither group needed to revisit an idea because of the fact that the teacher hinted strongly at (and sometimes even explicitly told) students the "correct" cycle idea.

Lack of idea development: Group 1. In general, students in group 1 were perfectly willing to record their initial ideas about a phenomenon in their Idea Journal. For example, the following table represents Darla's and Lacey's identical Idea Journal entries for Idea #2 in cycle 4. The question: "What happens to an object's energy during a resistive interaction?"

Table 6-1. Darla and Lacey's Idea Journal entries for idea #2, cycle 4.

Date	Statement/Drawing of Idea	Example or Evidence
1/20/00 (activity 2)	"There is less energy of an object in a resistive interaction."	"The sandpaper and wipe board. It takes more energy to push the object because of the resistance it loses energy."
1/21/00 (activity 3)	"No new entry."	
1/24/00 (activity 4)	"No new entry."	

The first entry was made at the end of activity 2 (What's a Little Friction?), in which groups pushed a wooden block across a wipe board and filled in various energy diagrams explaining the energy transfers in the system. At the end of the other relevant activities -- activity 3 (Slowing Down) and activity 4 (No Friction?) -- Darla and Lacey decided not to make any changes to the original idea.

Darla's and Lacey's entries follow a similar pattern for Idea #3 in cycle 4 (How does a resistive interaction affect motion?).

Table 6-2. Darla and Lacey's Idea Journal entries for idea #3, cycle 4.

Date	Statement/Drawing of Idea	Example or Evidence
1/21/00 (activity 2, 2nd day)	"A resistive interaction affects motion."	"A resistive interaction affects motion by making the object slow down."
1/21/00 (activity 3)	"No new entry."	

As their journal entries illustrate, it was not uncommon for group 1 students to write "the same", "no change", or "no new entry" for each idea revisitation after the first. In other words, Ideas 2 and 3 in this cycle (which are typical in this regard) underwent very little idea development over the course of the cycle.

Satisfaction with one's ideas: Lack of guidance.

One might suggest that group 1's lack of idea development in these Development activities could be due to a number of different things. Erroneous experimental results, perhaps, or a lack of mental effort.

Re-examining the intended process of idea development in Development activities, the original intent was that the experimental results from these activities would serve as foci of small-group discussions. In turn, these small-group discussions were to help students iron out the imprecisions or imperfections in the group ideas. So perhaps the trouble is with the results -- that the experimental results turned out differently than the developers intended? For group 1, this certainly was not the case. Group 1, an extremely conscientious group, performed the experiments as best as they could, and therefore typically got the results that they were expected to receive. So the answer must lie in the second portion of this process: the small-group discussion focused on the experimental results.

Recognizing that the discussion itself is somehow at the root of the problem, another possible explanation for the lack of idea development might have been that group 1 typically didn't put much effort into their sense-making discussions. As discussed in Result 7, below, this simply isn't true. Group 1 nearly always made the effort to complete the activities as intended and understand the ideas associated with those activities. Therefore, the issue wasn't with the effort associated with the small-group discussion, but with the discussion itself -- or, more precisely, the guidance contained within that discussion. But before I explain group 1's lack of idea development in terms of guidance, let me first situate the specific role of small-group guidance within the general utility and function of guidance (all types) in the idea development process in CIPS.

Guidance and idea development. As argued in chapter 2, there are tasks that students can do by themselves (reflecting their actual level of development) and there are tasks that students can do with guidance (reflecting their potential level of development). The process whereby a student's potential development level becomes his or her actual developmental level (i.e., the process whereby students learn to do for themselves what they could originally only do with help) is by internalizing guidance from others. In CIPS, the "tasks" that are most relevant are those instances when students must decide whether to accept, reject, or modify an idea -- and then, if modification is necessary, there is the further task of developing the idea appropriately. Guidance from the teacher and one's peers is often crucial to this task -- a task that appears in every activity in the CIPS curriculum cycle. Precisely how this guidance is relevant is outlined in more detail below.

For any given cycle, the assumption (and hope) is that group and class ideas will evolve in the following manner:

- Each group develops a pool of initial ideas about a phenomenon. Once it is agreed (through group discussion) that one or more of these initial ideas might be better than the others, the "better" ideas are shared with the class. The class now has a list of initial ideas to test experimentally. (Elicitation activity)
- Groups engage in experiments and discuss the effects of the experimental results on their group ideas. Consequently, groups discard ideas from the initial pool (perhaps leaving one or two), modify the remaining ideas, or perhaps even construct new ideas. These modified/new ideas may or may not be shared with the class. (Development activities)
- Each group makes final modifications to their group ideas, and then these finalized ideas are shared with the class. The teacher then leads an interactive whole-class discussion, in which the class agrees which ideas best explain the cycle phenomena. Ideally, these best ideas are very similar to the ideas accepted by the scientific community. (Consensus activity)

Clearly, guidance is a vital component of each step of this process. In elicitation activities, a student receives guidance as to whether his or her initial idea should be accepted, modified, or rejected as the group decides which set of initial ideas best explains the elicitation phenomenon. This guidance obviously comes from the other members of the group. The guidance in consensus activities is an inter-group guidance that (with some teacher assistance) helps the whole class decide (finally) which ideas really are best able to explain the cycle phenomena. Development activities, which are the mid-cycle activities most relevant to the lack of idea development considered in this section, consist almost entirely of groupwork; rarely, if ever, were whole-class discussions found in these activities. Therefore, in these activities, the guidance for the modification of ideas is almost wholly centered on intra-group guidance, much more so than in the elicitation or consensus activities.

Is the issue of guidance useful in explaining the early satisfaction with one's ideas, as illustrated in the journal entries of group 2? Let us explore this possibility. The key question: What happens when a group considers an idea that is incorrect or imprecise?

In considering an imprecise or incorrect idea, one possibility is that group members realize that there is something "wrong" with the idea. In many cases, when this occurs, students are able to provide the intra-group guidance necessary to help the idea progress. If the group is a motivated and well-functioning group, the group members can pose questions, construct arguments, and provide examples that should guide each other into reconsidering and redeveloping the idea.

There will also be instances when a group member is unhappy with his or her understanding of an idea, but no one in the group can provide the necessary guidance to help that person make progress. That is, there will be times when intra-group guidance is not sufficient to clear up the imprecisions that are recognized by one or more of the group members. The students might be talking past each other (i.e., they may each have separate ideas that aren't understood by the others), for instance, or one student may be simply unconvinced by another student's argument. In these cases, the students might simply agree to disagree, or they might also turn to external guidance (e.g., the teacher) to help them clarify or resolve their differences.

Yet, as a group considers an idea that is imprecise or incorrect, there is one other possibility. Above, I discuss cases where someone in the group realizes that the group idea is just that: incorrect or imprecise. But what happens when no one in a group recognizes that an incorrect or imprecise idea is problematic? More specifically, what if a group agrees that an imprecise or incorrect idea is the "best" explanatory idea for the cycle phenomena thus far, even if there is a better explanation in light of the latest experimental evidence? It turns out that it is this case that is most relevant to group 1's lack of idea development in this study.

When an imprecise or incorrect idea is up for scrutiny and no one in the group recognizes that there is something "wrong" with the idea, the typical result in CIPS is that the group members will record that idea in their Idea

Journal as the idea that best explains the cycle phenomena up to that point. Usually, this is both acceptable and expected; in CIPS, the plan is that students will be content with their ideas at various stages in the idea development process, and that new experimental evidence will help students decide which ideas are worth keeping and which ideas are in need of further development. But what happens when a group continues to put faith in an imprecise or incorrect idea, even when the group has been confronted with numerous experimental results that were intended to prompt further development of that idea?

The simple fact of the matter is that there are times when even interested and motivated students can't provide the guidance that is needed to help each other's ideas develop to the next level. In these instances, the underlying problem is that, based on the group members' current level of development, the group has reached its limit for internal guidance.

An example of this from group 1 is with Idea #2 from cycle 4: What happens to an object's energy during a resistive interaction? As seen in the journal entries above, the group started with a reasonable initial idea for the relationship between resistive interactions and energy ("There is less energy of an object during a resistive interaction."), but then failed to develop the idea further after additional investigation ("No new entry."). My argument is that the failure to develop Idea #2 in cycle 4 was based on group 1's inability to provide internal guidance during these later activities.

Unlike the later attempts at re-addressing idea #2, group 1's initial development of the idea is an example of successful internal guidance within the group. The idea was formed at the end of activity 2 from cycle 4 (What's a Little Friction?), where the group looked at cases where a wooden block sped up (forces involved: external push, friction) and moved at constant speed (forces involved: external push, friction). It was the following conversation that prompted the first written version of Idea #2:

L (reads) " What happens to an object's energy during a resistive interaction?"

D Um, the energy…

G The object goes slower.

L The resistance causes the object…

G …to go slower.

D Well, the energy kinda drains, for lack of a better word. Like, [the energy] get[s] less -- because there's resistance and pulling.

L Ok, so we have to put that in words.

D Ok, how about this? "There is less energy of an object during a resistive interaction." (D, L, G record this answer)

Here, group members were able to provide the guidance necessary for the group to construct and record its initial idea about resistive interactions and motion. Grace first offered a motion-based answer, but then Darla offered guidance in the form of reintroducing the idea of "energy" and offering her perception of energy as "draining". Then Lacey stepped in with some guiding thoughts of her own, which was to imply that the group's answer needed rethinking and rephrasing ("Ok, so we have to put that into words.").

Having looked at a case where guidance was successful in developing an idea, let us move on to the activities where Idea #2 was re-addressed -- activities, where, unfortunately, the idea was not developed any further, even though the original idea was imprecise in a number of ways.

In the next activity, activity 4, the group considered the additional case of a wooden block speeding up during a push, and then slowing down. The group accurately filled out the energy diagrams for the speeding up and slowing down portions of motion, and also held discussions that demonstrated their understanding that the object gained motion energy during the push (even though friction was acting) and lost motion energy after the push (once the person let the block go). At this point, then, the group had considered (and understood, seemingly) three phenomena related to resistive interactions: speeding up, constant speed, and slowing down. Therefore, ideally, an idea at the end of activity 4 should have been able to explain all three phenomena.

But did their idea change after activity 4? No. In fact, the conversation that prompted the group to leave their initial answer for Idea #2 unchanged was brief and rather superficial:

L [Activity 4] says that we have to do [Ideas] number one, two, and three.

D Again?

L Yeah.

D But they're the same. We still think the same.

.

.

.

L Do we have to do all [three ideas]?

D I just wrote "No new entry." (L writes this in her journal)

One must first recognize that there was definite room for development after the initial construction of Idea #2. Darla's initial idea was on the right track, but she was primarily concerned with comparing phenomena where resistive interactions are present to phenomena where resistive phenomena are completely absent. This would explain why her answer did not address the separate cases of the object speeding up and the object moving at constant speed, but instead tried to place all resistive-based phenomena under a single umbrella ("There is less energy of an object in a resistive interaction".). Her idea also did not specify that the energy of interest in these cases is <u>motion</u> energy. Had Darla considered both cases (speeding up/constant) and included the idea of motion energy, her idea might have taken this form at the end of activity 2: "When both a push/pull and a resistive force act on an object, the motion energy will either remain constant (if the push/pull equals the resistive force) or will steadily decrease (if the push/pull is less than the resistive force)."

At the end of activity 4, after the additional "slowing down" phenomenon was introduced, the idea could have developed even further:

> When a resistive force acts alone on an object, that object steadily loses its motion energy. When both a push/pull and a resistive force act on an object, the motion energy will either remain constant (if the push/pull equal the resistive force) or will steadily decrease (if the push/pull is less than the resistive force).

I'm certainly not arguing that Darla and Lacey should have developed this expert idea after activity 3, the first activity dealing with this idea -- or even necessarily developed this idea at the end of activity 4, which was the second time that this idea was addressed. According to the general CIPS plan, the cycle 4 consensus activity was ultimately to be the place where students might agree that this idea (or something like it) was the idea that best explained the cycle phenomena. On the other hand, no one would deny that group ideas are expected to evolve as the group performs more and more Development experiments -- even if the group never reaches the target ("expert") idea. In other words, I'm merely pointing out that Darla's and Lacey's idea (ideally) should have more closely approached this "expert" idea as the cycle wore on -- in spirit, if not in words; however, this did not occur. Darla and Lacey accepted their initial imprecise idea as essentially correct even after performing two additional experiments related to resistive interactions.

To repeat, my argument as to why group 1's cycle 4 ideas did not develop further was a lack of internal guidance within the group. One type of guidance that the group seemed unable to provide was the ability to point out that Idea #2 (relating to resistive motions and energy) should have been able to explain all of the resistance-related phenomena up to that point, and that it wasn't enough for Idea #2 to vaguely compare only resistive and non-resistive phenomena. In other words, no one in the group was able to suggest that the group should look back over all relevant activities and determine if their initial idea still explains the phenomena in those activities. Furthermore, the group seemed satisfied with any idea that somehow included "energy loss" -- since we all know, from experience, that friction typically results in some kind of loss (a loss of speed, typically).

Probing Darla's thinking a bit, it is interesting to note that it would be essentially correct to say "there is less energy of an object" if, by that, one means that a portion of the energy input is lost to friction, and so there is less energy available for motion energy. But it is not at all clear that this is what the group was thinking. And even if it was what Darla was thinking, this idea was never made explicit to any of the other group members -- meaning that, for group 1, the use of the term "energy" in Idea #2 was fuzzy at best. This is another reason why the group idea for resistive interactions and energy stalled after the initial idea. It appears that the group was unable to differentiate between the object's energy (motion energy) with the other energies in the system (e.g., the input energy) because of their level of understanding of energy concepts. Since the group could not differentiate between the two, it clearly was not possible for the group members to provide guidance that would direct the group to explore the finer workings of energy transfer within the system.

A summary of the above argument is that, for group 1 in cycle 4, its members were not at a level of development that could guide the group through an understanding of the details of energy and resistive interactions, and so the group bottomed out at an imprecise understanding that could not be changed without external guidance. And since this external guidance did not appear (from the teacher, for example, or from another student outside the group), the group became satisfied with their ideas after only one or two activities when, in fact, they were supposed to make continuous modifications to those ideas after each activity in a six-activity cycle.

A hypothesis related to internal guidance -- a hypothesis that needs further confirmation, but one that appears perfectly reasonable based on the above evidence -- is that groups are far more likely to reach their internal guidance limit when the group's current ideas are somewhat similar to the ideas that best explain the cycle phenomena. I would suggest that, if a group idea is extremely far removed from the "correct" idea (for instance, if group 1 had said that friction causes an object's energy to be more rather than less), there is a high probability that any new experimental evidence will cause at least one group member to be dissatisfied with the group ideas. But, as in the examples above, it grows easier and easier for a group to reach its limit for internal guidance when the group ideas are only "imprecise" or "unclear", rather than being completely off track. That is, the lack of clarity and lack of concept differentiation in the group ideas may go unnoticed because the group ideas are still vaguely correct, the result of which is that the group sticks with their previous idea, even if it is in need of some development. And if a group's idea is fairly reasonable after the very first development activity in a cycle (arguably, this occurred in Idea #2 for group 1, above), it is unlikely that the group will modify its initial idea over time, even when changes in the idea would be justified based on the experimental results from the activities later in the cycle.

Finally, to place my guidance-based argument in the context of the research that has previously been done on sense-making, there is clearly a connection between my argument -- that group 1's lack of idea development is due to the limit of their intra-group guidance -- and Hatano's claim that one of the basic requirements for sense-making is that the learner must be capable of recognizing the inadequacy of his or her comprehension. Scaling up Hatano's claim to encompass small-group activity, the requirement for group sense-making would seem to be that, if the group is in agreement on an idea, at least one member of that group must recognize the inadequacy of that idea for it to undergo further conceptual development. When a group member does finally recognize the faults of group idea, the "recognizing" group member is then in a position to bring these faults to the attention of the other group members; hopefully, this unearthing of the idea's problematic aspects then kicks off a significant amount of intra-group guidance -- guidance that, ideally, helps the group reconceptualize and reconstruct the group idea in such a way that it addresses the faults that were previously identified.

Lack of idea development: Group 2. Whereas group 1's Idea Journal entries often took the form of "No new entry" or "The same", group 2's entries tended to be nonexistent, at least after the initial entry for any given idea.

Sabrina left one of the seven ideas from cycles 3-5 completely blank. She failed to revisit three of the remaining six ideas. Roxanne left three of the seven ideas completely blank; of the four remaining ideas, Roxanne failed to revisit three. Arthur left one of the seven ideas completely blank; of the six remaining ideas, Arthur failed to revisit three. Also, regarding the entries that they did make, Sabrina and Arthur each failed to provide evidence for three idea entries, and Roxanne failed to provide evidence for one idea entry.

Here are sample Idea journal entries for Roxanne, Arthur, and Sabrina.

Table 6-3. Roxanne's Idea Journal entries for idea #1, cycle 3: Can pushes and pulls be combined?

Date	**Statement/Drawing of Idea**	**Example or Evidence**
1/12/00 (activity 1)	"Yes because you can combine different strengths, it make a stronger force."	"When you weight different objects individually and then combine the weight, it is the same as the total weight."
(activity 2)	(Blank -- no entry)	

Table 6-4. Arthur's Idea Journal entries for idea #2, cycle 4: What happens to an object's energy during a resistive interaction?

Date	**Statement/Drawing of Idea**	**Example or Evidence**
1/20/00 (activity 2)	"It starts to decrease."	"Push a block without friction and it goes on forever, push a block with friction, it will stop eventually."
(activity 3)	(Blank -- no entry)	
(activity 4)	(Blank -- no entry)	

Table 6-5. Arthur's Idea Journal entries for idea #3, cycle 4: How does a resistive interaction affect motion?

Date	**Statement/Drawing of Idea**	**Example or Evidence**
1/20/00 (activity 1)	"Same as Idea #2."	"Same as Idea #2."
(activity 2)	(Blank -- no entry)	
(activity 3)	(Blank -- no entry)	

Table 6-6. Sabrina's Idea Journal entries for idea #3, cycle 4: How does a resistive interaction affect motion?

Date	Statement/Drawing of Idea	Example or Evidence
1/20/00 (activity 1)	"The motion will not be fast."	"The resisting block makes the motion slow down."
(activity 2)	(Blank -- no entry)	
(activity 3)	(Blank -- no entry)	

Table 6-7. Sabrina's Idea Journal entries for idea #2, cycle 5: What kind of motion do falling objects have?

Date	Statement/Drawing of Idea	Example or Evidence
1/31/00 (activity 1)	"Falling objects have constant speed."	(No example or evidence given)
2/7/00 (activity 3)	"Falling objects are speeding up."	(No example or evidence given)

As seen above, the pattern of Idea Journal entries differed substantially from group 1 (Darla, Lacey, Grace, and sometimes Porter) to group 2 (Arthur, Roxanne, Sabrina, and sometimes Jasper). Typically, group 1 would spend some time thinking about their ideas, and then ultimately would decide that their ideas were in little need of change. This resulting "Idea Journal" pattern was that group 1 would record an initial entry, and then that entry would not change from activity to activity. Group 2, on the other hand, skipped many of the idea revisitations, and even skipped some of the cycle 3-5 ideas in their entirety. In these instances, the members of group 2 clearly did not put in the time necessary to decide whether their ideas needed modification.

Satisfaction with one's ideas: Not valuing the ongoing self-examination of ideas.

Result 7, below, outlines a number of differences between groups 1 and 2. Group 2 was off-task much more frequently than group 1, for example. Also, group 1 was fairly polite and cooperative, while group 2 tended to be less polite and somewhat individualistic. Finally, and perhaps most importantly, group 1 made an effort to complete the activities as intended and understand the activities' underlying ideas, while group 2 tended to follow its own goals, some of which (at different times) included socialization, filling in the worksheet blanks with some sort of response, and playing with the experimental equipment -- although, to be fair, there were also times when group 2 made a real effort to understand the activity/cycle ideas, although these occurrences tended to be somewhat rare.

The blank spaces and lack of follow-up in group 2's Idea Journals follow logically from the group's basic characteristics, as described above and in Result 7. If the students in group 2 were switched back and forth between filling in blanks (which meant that sometimes group 2 wasn't terribly worried about ensuring that the initial group ideas were accurate and comprehensive) and socializing or going off-task (which meant that the group didn't care whether they had an idea written down at all), there's certainly no reason to think that the group would be motivated to revisit their initial ideas in order to test their validity and make them more accurate and more comprehensive. To put things another way, group 2 (for a variety of reasons) did not value the ongoing self-examination of their physics ideas -- regardless of whether or not these ideas needed changing. Perhaps the best way to succinctly summarize group 2's attitude toward the Idea Journals is to quote Arthur, one of the group's members: "I hate these stupid things."

An interesting by-product of group' 2 s attitude toward idea development is that, because the group often did not put as much effort as was necessary into constructing their ideas early in the cycle, there was at least one consensus activity where the group was forced to put a good deal of effort into writing and discussing their ideas during the activity's "Prepare your wipe board!" sub-section. This accounts for the record-high percentage of 16% for the percentage of time that group 2 engaged in SMD during the "Prepare your wipe board!" sub-section in activity 5 from cycle 5 (Putting it All Together, the consensus activity). (Not even group 1 had a percentage this high for a "Prepare!" sub-section.) Had this end-of-cycle effort by group 2 been more common, Result 3 (6% of time engaged in SMD in "Prepare!" sub-sections) might have been different. However, the effort that group 2 spent on reconsidering and modifying their group ideas (which was not exemplary) tended to carry over into the small-group consensus discussions as well.

As related to Hatano's requirements for sense-making, which were presented in Chapter 2, it is fair to say that the members of group 2 often did not meet one of the sense-making requirements: seeing comprehension of the target rule, procedure, or concept as important. As seen in group 2's failure to effortfully revisit their ideas, this requirement seems to be even more important in CIPS than in traditional curricula because of the fact that sense-making in CIPS is extended over the course of an entire cycle, and so the viewing of comprehension as important is something that must be maintained for days at a time. In the case of group 2, there were many times where students didn't consider comprehension to be important the first time that they addressed a cycle idea, and so there was little chance that they group would suddenly adopt a "comprehension is important" attitude in subsequent reconsiderations of the same idea -- although, as described above, there was one consensus activity in which this occurred.

Satisfaction with one's ideas: Teacher hinting or telling. The basic premise of the CIPS curriculum is that students are to be largely responsible for the introduction, development, and critical evaluation of ideas, models, and explanations. But what of the teacher's role in this student-centered process? If the students are to do most of the work, where does the teacher fit in?

In CIPS, the teacher's role is not to hand out correct answers or make heavy-handed judgments about which student ideas are the most "correct". Instead, the teacher plays the general role of "guide", which includes the responsibilities of ensuring that all students share their ideas, that they do so in a polite but critical manner, that the scientific experiments run smoothly, and that students carefully consider the results of the experiments as they reflect on and reconceptualize their physics ideas.

Clearly, it would be nice for CIPS teachers to always stick to this role of "guide", a role of much less authority than the traditional teacher role of lecturer and supplier of answers. In this study, though, there were times when the teacher strayed from his suggested "guide" role in that he sometimes hinted strongly at (or even explicitly told) the students which ideas/evidence were correct. The result of these tellings and hintings is that students knew that the hinted/told ideas were correct, and so there was no need for the group to revisit and rediscuss the ideas in later activities. Quantitatively, this effect translated into a smaller percentage than expected for the "Prepare your wipe board!" sub-sections in the consensus activities, the result that I'm explaining in this section.

One such example occurred when the teacher effectively pointed out the best answer to Idea #1 in cycle 3 (Can pushes and pulls be combined?) and then provided his own wording for the idea, with evidence. This occurred just after the presentations in the cycle 3 elicitation activity (Can You Lend Me a Hand?), where the only remaining task was for the students to fill out the Idea Journal for Idea #1. (In the elicitation activity, groups were asked to weigh two separate objects and then predict the combined weight of both objects.) The following excerpt shows how the teacher identified a "best" answer (which happened to belong to group 1) and then told the students exactly what to write in their idea journals.

T You just completed activity 1. The statement at the top of [the Idea Journal] is the statement [Idea #1] that we want to address. In fact, you already did that. If you remember our presentations that we did yesterday…the presentations all had to do with this question: "Can pushes and pulls be combined?"

T (to group 1) What about you guys? The reason I call on you is because your team came up with a pretty good [answer] yesterday. Do you remember what it was?

D (flips through pages, then reads) We put: "Pushes and pulls can be combined. When they are combined, the strength of the pull becomes larger than if one object was pulling." Our evidence: "From [the experiment], when you add object #1 and object #2 you would get their pull combined."

T Ok. Let's see if we can put it into words simply. "Can pulls and pushes be combined?" Yes, but we need more specific evidence. Can we say something like: "They can be combined because…if we weigh two individual objects and then weigh them combined, the weight combined equals the sum of the two individual weights."

As you might have expected, the entire class dutifully copied this answer into their Idea Journals -- meaning that the students were now sure to have the "correct" answer, and so they had no real reason to reconsider this idea at a later date. The idea was originally meant to be re-addressed in the very next activity (activity 2, Lots of Pulls!), but -- after the teacher's rewording and dictation -- there wasn't any need for further student development of this idea.

In cycles 3-5, there were other instances of the teacher hinting strongly at (or explicitly telling) the class the correct way to think or reason about something. One example was when the teacher introduced the cycle 4 elicitation activity (Will It Slow Down?) by saying that "friction has now entered the world" (in both classes). From the researcher's perspective, this was an odd thing to do, since the whole point of the cycle 4 elicitation was to figure out why objects normally slow down on earth.

Another example was when the teacher made the following off-hand comment (in class 1) while introducing the "atmosphere" portion of activity 2, cycle 5 (Exploring What Causes Gravity):

T [In the elicitation activity] you came up with the idea that the air -- the atmosphere above us -- causes gravity. And that kinda makes sense, because if you start climbing up a really tall mountain, the further up you go, the less you weigh -- and the less air there is above you.

When students were later asked in activity 2 to predict what would happen if a mass was put in a jar, and then the air was evacuated from the jar, Darla wrote that "The mass [weight] would become less, because when you go up in the mountains you weigh less". In the previous activity, Darla had already written that she thought that the earth's atmosphere was the cause of gravity, but the teacher's statement provided Darla with additional evidence to back up this idea. The precise result of the teacher's words was that Darla appropriated the teacher's reasoning about air and gravity and made it her own.

Another class-level example of teacher telling/hinting was when the teacher was forced to engage the class in a heavily guided discussion after the class had just watched a video on the Cavendish experiment (from activity 3 in cycle 5, More Exploration of What Causes Gravity)[5]. In the discussion, because the video wasn't as clear as it should have been, the teacher ended up explicitly telling the classes that gravity is based on mass (not size), and that our bodies pull on the earth just as much as the earth pulls on us. He said the latter to refute the student idea that gravity is limited to big objects being attracted to small objects.

One last example of hinting/telling, this time limited to a conversation with a particular group, was when the teacher told group 1 that a held ball has a force holding it up. Before the teacher arrived, Darla had been convinced that held objects do not have forces acting on them (a common alternative conception). After the teacher's arrival and comment, Darla had no choice but to accept the teacher's answer as valid.

Overall, I don't want to give the impression that the teacher strayed too far from his "guide" role and engaged in widespread handing out of correct answers. This certainly was not the case. There were many cases where students asked the teacher to refute or validate their ideas, for example, and he refused to do so. On the other hand, the few times that the teacher strayed from his role did have an impact on group sense-making discussion. That

[5] The Cavendish experiment showed that bottles of water are attracted to boxes of sand, thus

result was that students occasionally knew the "correct" explanation for a particular phenomenon, and so there was no need to reconsider that idea at a later date. It was this circumventing of the cycle-long restructuring process that contributed to Result 3: the percentage of time that groups engaged in sense-making was lower than expected in "Prepare your wipe board!" sub-sections.

Time is short. At this point, we've come to realize that there were a number of factors that contributed to the low percentage in "Prepare your wipe board!" sections in the consensus activities: lack of intra-group guidance, not valuing the day-to-day development of ideas, and the explicit or implicit teacher validation of certain cycle ideas. And, although I hate to put the final hole in a sinking ship, there is one additional factor that might have contributed to the low "Prepare!" percentage (6%). Even if students had wished to engage in SMD in "Prepare your wipe board!" sub-sections -- which would be a doubtful assumption for either group, for the reasons described above -- there was the issue that, in wanting to complete the consensus activity in one period, the teacher always told groups that their presentation boards should be up in 5-10 minutes time. Knowing that students necessarily had to spend a significant portion of the time simply writing their answers on the presentation boards, this time restriction almost ensured that there would not be much time available for the revisiting and restructuring of the group ideas.

---

helping to prove the idea that all objects exert a gravitational force on one another.

Result 4: Sense-making Differences between Individual Sub-sections

Result 4: The percentage of time dedicated to sense-making discussion in individual sub-sections varied quite a bit. Some sub-sections where sense-making discussion had been expected contained very little to no sense-making discussion.

Rather than account for the wide variation of percentages from sub-section to sub-section, I will only provide an explanation for the sense-making percentages in the sub-sections with the top four highest percentages and bottom four lowest percentages (averaged over both groups). In this manner, we might learn something from the "best" and "worst" sub-sections in terms of how their successes might be repeated (and their failures avoided) in future curriculum design projects.

Given that the CIPS curriculum was organized such that most small-group sense-making discussion would occur in the "Making sense" and "What does your team think?" sub-sections, it comes as no surprise that the top three sub-sections were of these types. These sub-sections are still quite notable, however, in that the top 2 "Making sense" sub-sections achieved twice as much SMD as the average "Making sense" sub-section (an average of 20%, from Table 5-16), and the top "What does your team think?" sub-section (at 37.5%) was also far above average for its type (an average of 25%, from Table 5-16).

"Making sense", from activity 4 in cycle 4 (44%). Activity 4 (No Friction?) began with a whole-class discussion of sports that focused on which sports might be playable without friction. Next, the students viewed a Magic School Bus videotape (a cartoon) depicting what motion on a baseball field would be like without friction: freely sliding players, changes in direction due to thrown or caught objects, and so forth. After watching the video, students were then asked in "Making sense" to name reasons why they would and would not want to live in a frictionless world. Discussion excerpts related to these "Making sense" questions are provided below.

Group 1:

L	I don't know why we'd want a frictionless world. I get thrown up in the air and I stay there. That's fun, huh?
G	Would you ever want to stay there forever?
L	No -- I know. The only sport I do is cheer (cheerleading) and I get thrown up in the air and am going to stay there and float around.
G (to L)	Then you go by a tree and pull yourself down.
D	You have to use energy. All you have to do is push off a wall and you're on your way forever.
L	You won't get tired. You can run the mile no problem.
D	You could slide the mile, not run it.
G	You'd fall.
D	If you tried to run it, you'd fall every other step.
G	You'd better know how to skate.

Group 2:

R I think it'd be fun to live in a frictionless world.

S Not me, because I wouldn't want to be running into everything.

A The only way I'd like to live in a frictionless world is if it was only the ground that was frictionless.

A I don't want the walls and tables to be frictionless. That would be a pain in the butt.

R Our olympics would be just, you know...(mimes sliding)

A If you're sitting there trying to eat, and you set a cup down on the table, if everything's frictionless...when you set the cup down...(mimes eating and grabbing cup) ...the cup would keep trying to go away with you.

R It depends if there's a slope or something. Unless you drop it a certain way.

A Humans can't make something...human hands can't set something down and keep it from moving. It would always be moving.

S I still think ice skating [can be done without friction].

A Ice skating you have to push off.

A I was thinking about swimming during the movie, but then I decided "no". (can't do without friction)

S You'll die before you can even keep going.

A Why?

S You'll drown.

A What makes you say that? Why would you drown in the pool.

S Because you'd be like moving, and moving, and then you'd drown.

A But the water would still be the same.

S If you were underwater, you'd run out of breath!

A (to S) Yeah, you're right. That wouldn't be good.

What you can't see in these discussion excerpts is the amount of smiling, laughing, and generally positive emotional response that accompanied these sense-making discussions. This is one of the unique characteristics of this discussion, in fact. There were many instances in CIPS where students were laughing and smiling, but those instances tended to be associated with off-task behavior or the setting up and running of hands-on experiments. This "Making sense" sub-section is one of the few sections where students enjoyed the sense-making discussion itself. This alone is enough to mark this sub-section as extremely successful in comparison with the other sub-sections in this study.

The reason for the interest and enthusiasm in these discussions is that, overall, the activity was a nice example of an intrinsically motivating activity. In chapter 2, the characteristics of intrinsically motivating activities were listed as: personal relevance, task novelty, and the degree over which the learner has control over the activity. Here, personal relevance was first established by the whole-class discussion around sports and friction, which was then maintained in the "frictionless baseball" section of the Magic School Bus Cartoon. The "Making sense" questions than provided a creative avenue for the students to explore how a lack of friction might affect their own lives (again maintaining personal relevance). These open-ended questions (e.g., Why would you want to live in a frictionless world?) gave the students the freedom to be freely imaginative, which is one way to ensure the second characteristic of an intrinsically motivating activity: that the learner has control over what occurs in the activity. Finally, the activity also happened to be novel in that videotapes (especially cartoons!) were fairly rare, and never as off-the-wall as this "frictionless baseball" episode.

"Making sense", from activity 3 in cycle 5 (41%). The high percentage of sense-making in "Making sense" in the More Exploration of What Causes Gravity activity (activity 3, cycle 5) is almost exclusively due to the extremely high percentage (62%) in group 1, which happens to be the highest percentage recorded for any sub-section in cycles 3-5 in the Force and Motion unit. The main factor behind group 1's extraordinarily high percentage is that Darla experienced an intense cognitive incongruity as a result of one of the "Making sense" questions.

Activity 3 in cycle 5 began with a videotape showing the classic Cavendish experiment, the experiment showing that bottles of water are attracted to boxes of sand, thus helping to prove the idea that all objects exert a gravitational force on one another. After it was clear that the video left most students confused, rather than enlightened, the teacher held a class discussion in which he made clear that all objects exert gravitational forces on each other: strands of hair, planets, and any other objects that you care to name. (Which was the point that the video had tried to get across, but had failed to do so.) After the class discussion, groups continued on to the next sub-section: "Making sense". This sub-section directed the students to place two pencils side-by-side on their desks and explain why the pencils don't attract one another.

At this point, Darla and the rest of her group held a brief discussion.

L Put two objects next to each other. (L, G put their pencils down beside each other.)

G What do we do with that?

L They didn't move closer.

G Why didn't they go next to each other?

L I don't know. Why don't they move together?

D We have to have a bigger object.

L We know they're not going to move together...unless you go like (moves table to get pencils moving).

G Cuz they don't have that kind of force in it. They're not mechanical. They're not robots or anything.

D Why do they not move? I don't get it.

G Cuz these just sit there. They don't have that kind of force to move. They weren't made to do that.

D If we hung them by a string [like in the Cavendish experiment] I bet they'd move.

experiment] I bet they'd move.

G Yeah, through the air.

L I put (reads): "No, they don't move towards each other. They don't have a force to push them together." (closes notebook)

At this point, although you can't tell from the transcript, Darla's facial expression and body language indicated that she was clearly unsatisfied with the answer -- although, since the class was ending at this point, Darla wasn't free to pursue the matter further. But Darla got another chance to clear up her confusion when, at the beginning of the next day's class period, the teacher told the class to complete "Making sense"; he did so because many of the groups hadn't started the section during the previous class period.

During this revisiting of "Making sense", group 1 held another conversation that picked up where the previous conversation had left off.

D I don't get this. Why don't [the pencils] go together? I don't understand that.

G What?

D The two objects.

G Like the water thing and the box?

D No. It says (reads): "Put two objects on your desk and watch them very carefully for 15 seconds. Do the objects move towards each other? Why or why not?"

L Uh huh, watch. (L rolls a pencil on the table)

G They're going to move against each other if you <u>push</u> them. It won't move [normally]. They're not supposed to do that.

(roll into each other)

D Why did they not move closer together? I don't get this.

L Because there's not a force of gravity on the side. There's a force of gravity in pulling it to the ground and holding it in one position.

D Well, yeah. But if you have the two water bottles - meter stick thing, that's on the string...well then the other one...

L Yeah, but they have a force drawing them to the sand.

G What's the force?

D The force is gravity.

L Yeah, right.

D But don't you think there's gravity between these two objects? (D points to two other objects: a pen and binder)

L Yeah, but it's not strong enough to pull the two things together. It's not like...it's different from if there's two pens hanging right here. (L holds out two pens side-by-side)

L And then this is sand (motions to an imaginary box) and this is sand (motions to another imaginary box). And then they (pen and box?) go towards each other because there's forces right here...pulling them to each other. And the two pens...the two water bottles...

G But they're not hanging. That's your only problem.

D (to L) So you're saying gravity only happens if there's hanging?

L No, gravity doesn't always happen if there's hanging. Because we aren't hanging right here [on earth].

D But, yeah...but how...

L Gravity's in the ground...like, gravity's not in the ground...(puts two pencils on table vertically, then lets go)...look, they don't go together. They fall...they may fall towards each other, but that's because pens aren't made to stand up on their end. They're made just to lay here. What I'm saying is: if these pens were attached to the ruler - how they proved gravity or whatever [with the Cavendish experiment], and there was a box of sand [on one side of the ruler] and a box of sand [on the other side of the ruler], and [the sand] pulls [the pen so that the stick rotates], they don't have a force that's pulling [the two pens] to each other.

G A table with no force. (laughs)

L There's nothing in between them that would be pulling them together. Just air.

D There's nothing in between the other ones, either. (the pen and notebook that D had just referred to)

L But there's nothing...there's nothing like sand or something pushing them together.

G I don't think there would be. I don't think sand...

L If there's something <u>in between</u> them (puts pen top between pens)...they go together.

G Well, they didn't move together! They just rolled because of the ground. (L pushes them together again) They're not moving! You're moving them.

D I still don't get it.

Judging from this conversation and the conversation from the previous day, it is quite clear that the pencil-pencil "experiment" gave rise to one of Hatano's basic conditions for sense-making: cognitive incongruity. In this instance, the incongruity came from the fact that Darla knew that gravity affects all objects, and so she simply could not understand why the pencils' gravitational forces did not move the pencils closer together. The result of this incongruity on group 1's small-group discussion was that, time and time again, Darla expressed her confusion about the experimental "results" and communicated her desire to construct an explanation that would clear up her confusion; in response, her groupmates did what they could to come up with alternative explanations and provide appropriate supportive evidence.

The unfortunate aspect of this particular "congnitive incongruity" is that the incongruity was completely idiosyncratic. Looking at my fieldnotes for the two days that the students spent on activity 3, it is fair to say that Darla was one of the few students to experience cognitive incongruity of this magnitude from the pencil-pencil "experiment". For example, group 2 was satisfied with its answer that "friction" caused the lack of motion, even though Arthur (who supplied this answer) admitted that his explanation was a complete guess.

Therefore, the factor that explains the extraordinarily high sense-making percentage in "Making sense" in activity 3 from cycle 5 is an isolated cognitive incongruity -- one that may or may not ever be repeatable in other students or other groups. However, supporting factors that allowed this incongruity to drive group 1's sense-making discussion were a) Darla's general interest in understanding the materials, and b) Darla's status of group leader within group 1. These two factors are described in detail below, in Results 7 and 8.

<u>"What does your team think?", from activity 1 in cycle 1 (37.5%)</u>. As I already mentioned, it comes as no surprise that the top three sub-sections in

terms of the percentage of time spent on SMD included a "What does your team think?" sub-section. "What does your team think?" sub-sections gave students their first opportunity to explain and predict in the context of a new, interesting phenomenon (in theory) -- and so a high sense-making percentage for these sub-sections is both reasonable and expected. Still, the fact remains that one particular "What does your team think?" section supported sense-making discussion far better than the other sections of its type. For this reason, the logical course of action is to identify the factors that allowed sense-making discussion to flourish in this sub-section relative to others of its type.

First stop: the obvious factor. The reason for the comparative failure of the cycle 4 "What does your team think?" sub-section has already been substantially documented in Result 2, the result dealing with level 1 and level 2 predictions in CIPS sense-making discussions. A key point of that analysis was that the predictions in the cycle 4 elicitation were just too easy. Almost down to the person, students knew that skateboards, shopping carts, and bicycles would eventually slow down if initially given some sort of horizontal motion. Therefore, no sense-making discussion was necessary for the elicitation; everyone knew what would happen (the objects would slow down), and everyone knew why (friction).

In contrast, the prediction question for "What does your team think?" in activity 1 from cycle 3 (Can You Lend Me a Hand?) was a nontrivial prediction, in which students were asked to answer the following question about an imagined tug-of-war: "What does your team think? Will it help Kinet's side if several of the aliens pull together against Teract, even though Teract is stronger than the others? What is your reasoning?" As seen below, group members made substantive predictions (with evidence) regarding this phenomenon.

Group 1:

L	I wrote that if they have two...if there's two against one, and the two strengths add up to equal to Teract's, then I think it would be a fair game. They would -- Stas and Kinet -- would have an equal chance of beating Teract, and Teract would have an equal chance of beating them.
D	Yeah. That's what I put, basically.
L	Because if their weights add up to... or if their strengths add up to the same amount of strength that Teract has, then they have a fair chance of beating Teract...because it's two against one, but it'd be a fair game because it's strength, not how many people they are.
D,P	Right.
D (to P)	What did you put?
G	Same thing.
D	...if Teract and Stas pulled, it would be...it would... ...both teams would have an equal chance.
L	If Stas and Teract...if Stas and Kinet were on one team against Teract...
D	It would be an equal game.
L	Yeah.

Group 2:

R I think it's because...let's say on a scale of 1 to 3, and Teract's like three...

A I said "just like math". Teract equals 4, Stas and Kinet each equal 2. You add the two smaller things together equals just as much as the big number.

S Why don't you do Teract equals 2, and Stas and Kinet equal 1?

A (to S) Same thing. Just smaller numbers.

J I just put "the more aliens, the more weight and strength".

A Yeah. I said the same thing. I just used a more complicated means of explaining it.

J That's pretty much what it is.

R The more aliens, the more power.

A Yeah.

S Well then if it's 3 to 1, then they'll definitely win. But if...

R Because the more strength.

A Yeah. Power or strength. Either one.

S So, what?

R Because the more strength, the more power. The more of an effect they would have.

In both cases, the groups were explicitly directed to make sense of a tug-of-war phenomenon -- and they did so. But only because the phenomenon was abstract and complex enough -- meaning that the precise amount of weights/strengths per side was not specified -- such that the general result of the tug-of-war (with supportive reasoning) was neither simple nor obvious.

A more interesting comparison is between the cycle 3 and cycle 5 "What does your team think?" sub-sections. What makes the difference in SMD between these sub-sections more challenging to explain is that the cycle 5 elicitation topic (the cause of gravity) was clearly not trivial. Experience has shown the reverse to be true, in fact. In observing elicitations in CIPS classrooms for the past 2 years, it has been noted that the topic of gravity -- a mysterious, invisible, "everyday" phenomenon -- consistently provokes one of the best whole-class elicitation discussions in the entire curriculum. This certainly was the case in this study, where class discussions about the possible causes of gravity were rich and substantive. In these discussions, students brought up a wide variety of possible causes (the earth's rotation on its axis, the earth's rotation around the sun, the earth's atmosphere, magnetism, friction), and often -- after some cross-student discussion -- provided a number of examples that supported or refuted the suggested causes for earth's gravity. In cycle 5, then, why weren't the small-group discussions that immediately preceded these class discussions equally as amazing?

The factors explaining the differences in the sense-making percentages between the cycle 3 and 5 "What does your team think?" sub-sections are specific to the group, as analysis will show. Group 1's percentage was lower in cycle 5 compared to cycle 3 because the group spent most of their time reading their ideas aloud (an activity not considered to be sense-making discussion). The sense-making that did occur was extremely brief because Darla's idea was almost immediately recognized as being the "best"; it was recognized as such because Darla was the only group member to offer reasonable supporting evidence for her idea. Group 2's percentage was lower in cycle 5 compared to cycle 3 because the group chose this particular activity as a time to be off-task, which meant, in this instance, that group 2 didn't choose to take the activity and its associated small-group discussion very seriously.

The following is an excerpt from group 1's discussion on the possible causes of gravity:

G [The cause of gravity is that] there's gravity on earth that lets things fall and not, like, float around.

D Right...so you...basically you said that gravity pulled it to the ground.

G,L Yeah.

L I said that gravity, like, has something to do with the atmosphere, or something. Wanna hear? Mine is...I think gravity is...um, something in the earth's air that...the earth's air has magnets. Like little magnetism things in it that pulls things to the ground. (shrugs)

D If they're little magnets, wouldn't they be in the earth to pull things down?

L (makes a face)

G Maybe it's only the ground that makes things go down. (laughs)

L Maybe there's something in the core...at the core. You know how no one can ever get past the...outer surface, you know? Maybe there's something in the core that pulls everything into...onto the ground.

D (to G) What do you think?

G I put that there's a force around the earth that makes things go down. Like, it keeps all the air inside, so it could fall...

D The atmosphere?

G Yeah, the atmosphere. It just sits staying there and just float around there. (L,D not really listening)

L (to D) Ok, what'd you put?

D I put that I thought gravity was caused by the atmosphere because...uh, once you go out of the earth's atmosphere, there's no gravity.

G Yeah, and makes things float around.

L I thought that gravity is in the atmosphere, right? Or has something...

(group writing)

L (writing) "We think that gravity..."

D (writing) "...is caused by the atmosphere. Because once you leave the earth's atmosphere…there's no gravity." (G, L record this answer as well)

As you can see, this discussion was generally limited to the listing of each group member's idea. After its initial "listing of ideas", the group implicitly settled on Darla's answer as reasonable because she was the only group member with evidence to back up her idea -- and probably also because, in the group's experience, Darla was the group member whose ideas typically proved to be correct. Grace's explanation for gravity was largely tautological ("There's a force around the earth that makes things go down") -- and so there wasn't really much to explore or question about her idea. Clearly, Lacey was not wedded to her idea that magnetism might be the cause of gravity, and so Lacey didn't feel the need to put much effort into defending this idea. When Darla questioned this idea, Lacey immediately backed off, shrugged, and wondered instead if there might be "something in the core" that causes gravity.

It is interesting that, by the end of the conversation (although you can't tell it from the transcript), the group hadn't even bothered to write down Grace's and Lacey's ideas. The group almost immediately accepted Darla's idea as the "group" idea, and likely would have never bothered to bring up the other ideas again. It just so happened that, a few minutes later, the teacher noticed that the group had only written down one idea, and made a comment to that effect. The teacher's comment prompted a very brief exchange in which Grace and Lacey re-iterated their ideas; afterwards, the group members added these ideas to the list of "group ideas" on their worksheets. No sense-making discussion was associated with this re-iteration and copying of ideas, however.

One matter of interest regarding group 1's lack of sense-making discussion in this small-group elicitation discussion is that, in part, the group's near-instantaneous adoption of Darla's idea was a matter of guidance. In a whole-class discussion, it is likely that at least one student could provide reasonable evidence (guidance) as to why magnetism might be a possible cause of gravity. This providing of evidence would keep the idea alive in the minds of the class as a possible cause, which would prevent the idea from being immediately tabled -- which is what occurred in group 1 when Lacey could not adequately support her idea. This capacity for better evidence-related guidance in whole-class discussions may partly explain why the class discussions on gravity tended to be more animated and in-depth than the small-group discussions on the same topic.

Group 2's "What does your team think?" experience in cycle 5 is almost not worth explaining, as the group only spent a minute or so on the topic. The half-serious answers that resulted from their conversation were that "God" and "magic" were the probable causes of gravity on earth. In short, the main factor behind group 2's low sense-making percentage is that Arthur and Roxanne were in an especially silly mood that day -- a mood that proved not to be very conducive to SMD.

Returning back to our original result, which was that the sense-making percentage in cycle 3's "What does your team think?" sub-section was quite large (especially in comparison with the same sub-section in other cycles), the factors that explain this result are that: a) "What does your team think?" sub-sections are naturally supportive of small-group sense-making discussion, b) the prediction in the cycle 3 section was not trivial, unlike the prediction in cycle 4, and c) in the equivalent cycle 5 section, only one group member (Darla) could provide reasonable evidence for her idea, and so her suggestion for the cause of gravity (the earth's atmosphere) was immediately adopted as the "group" idea.

"Sliding blocks", from activity 3 in cycle 4 (34.5%). As with some of the previous results, there are both general and group-specific factors that explain the large sense-making percentage in "Sliding blocks". The general factor is that explanatory activities that require students to construct energy and force diagrams tend to provoke in-depth sense-making more readily than explanatory activities that do not require these diagrams; this is due to the conceptual scaffolding provided by the CIPS energy diagrams and the demand for precision required by the quantitative (but not necessarily numerical) force diagrams and speed arrows. The group-specific factor is that group 2 was led through an in-depth clarifying discussion of speed because Roxanne atypically chose to use her group leadership to demand that the group help her deal with a cognitive incongruity related to speed.

Of the two groups, group 1 was the only group to construct the energy and force diagrams in this sub-section; group 2 spent so much time on their clarifying discussion that they never made it as far as the diagramming activities. However, there are many examples over cycles 3-5 that support the argument that energy and force diagrams seem to be especially conducive to group sense-making discussion, as argued below.

As one of the grand unifying concepts in physics, energy is an incredibly powerful, yet also incredibly abstract means of describing and analyzing the interconnected parts of a physical system. Because it is so abstract, initial student descriptions of systems in terms of energy tend to be amorphous, imprecise sentences in which students confuse (or ignore) the different types of energies and energy transfers, and also do not identify (or differentiate between) the various parts of the system (the pushed object, the air, the ground, etc.)

Evidence for students' trouble in applying the energy view of matter to physical systems has already been introduced in this study. We have seen group 1's struggle to correctly and precisely apply energy ideas to objects in activity 2 of cycle 4 (What's a Little Friction?). If you recall, in this activity, group 1's discussion of an object slowing down due to friction was basically limited to "There is less energy of an object during a resistive interaction". (Note that group 1 was one of the more careful and thoughtful groups in its classroom, and so we can be confident that other groups in this class experienced similar troubles.)

A relevant issue relating to the context of energy-based explanations is that, when students made vague explanations similar to "there is less energy of an object" (above), they were being asked to respond to an energy question using only words. No diagrams or drawings were involved in any way. Compare group 1's imprecise energy description of a slowing object, above, to the conversation that was held by the same group in the "Sliding blocks" sub-section in activity 3 from cycle 4 (Slowing Down), a conversation that was driven by the energy diagram shown in Figure 6-1.

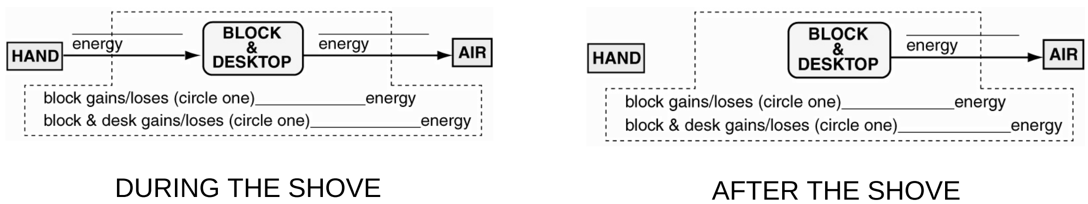

Figure 6-1: Energy diagrams in "Sliding blocks" section, from the cycle 4 Slowing Down activity

L (reads) "Fill in the chart on the next page to analyze what happens during the shove and while the block is just sliding." (to D) Would [the energy of the hand acting on the block/desktop] be mechanical?

D Yeah. (L writes)

L And then the [transfer of energy from the block/desktop to the air] would be heat energy, right? The block...does the block gain or lose [energy]?

D Loses motion energy -- because it slows down.

L And then gains thermal?

D Yeah, I think so.

D Oh, wait. This is during [the shove]. So it gains...motion energy.

D Then it's "lose". (the block loses energy after the shove)

L And then it's...so [the block] gains heat, right?

D [After the shove] it's "loses". [During the shove] it's "gains".

L It's still heat energy from the block, to air? (after the shove, compared to during the shove)

D Would the [energy transfer for the block/desk after the shove] be "loses"? But would it lose thermal energy, or would it...wait.

L Wouldn't it lose it, because... you pushed it and it's slowing down. And so the thermal is like...(pushes block) it's like...cooling down.

D Yeah.

Although the students in this excerpt still need to perfect their understanding of energy -- the table gains, not loses, thermal energy after the push, for example -- it is clear that the energy diagrams on these worksheets provided the scaffolding that Darla and Lacey needed to more adequately apply energy ideas to the hand/block/air system. Rather than limiting their explanation to something along the lines of "the block had less/more energy", Darla and Lacey were able to differentiate between the different types of energy (motion energy, thermal energy) as well as differentiate between the components of the system (the hand, the block, the air). Initially, students do not make these important distinctions in their energy analyses, which is why the scaffolding in these energy diagrams is crucial in helping students to clarify the many facets of energy analysis in their own minds -- as well as in the minds of their fellow group members. In this case, the diagrams' scaffolding of the students' energy analysis was a key factor in this group spending a good deal of time on sense-making discussion, the result that is being explained in this section.

Another activity that was successful in supporting SMD in the "Sliding blocks" sub-section was the drawing of force diagrams and speed arrows. Consider the following conversation from group 1, which was centered on drawing force diagrams and speed arrows for the same phenomena described above: the motion of a pushed block during and after the push.

L (reads) "Draw and label the arrows representing the pushes and pulls on the block." Let's see: [during the push] the hand pushes...and then the friction...

D Is less than the push.

L Is it the same thing [after the push], but that friction is just...greater?

D Yeah, but there's no longer a hand pushing it. So I think it's just friction. Because it's already sliding.

L Ok. (reads) "Draw speed arrows to represent the block's motion." So, it's...so...it kinda speeds up, and then slows down.

D Yeah. It speeds up when you're shoving it, but when you're not it slows down.

L So, wouldn't it speed up and slow down?

D Yeah, but it wouldn't speed up very much.

L So, wouldn't [the block's speed] go, like...up [during the push], maybe, and then down [after the push]?

push], maybe, and then down [after the push]?

D Yeah.

It's one thing to be vaguely aware that there are two forces acting on a pushed block (the external push, friction), which is the description that a student might give if only asked to provide a purely qualitative (i.e., conceptual, non-representational) analysis of the pushed block before and after the shove. Where the drawing activity was successful in pushing SMD beyond this basic level was that, by the very nature of the exercise, students were prompted by the representational, quantitative[6] nature of the activity to go beyond a two-force description of the phenomenon and consider a number of additional conceptual issues:

Is there a force arrow in the direction of motion after the push, or just while the hand is in contact with the block?

Are the lengths of the push arrows the same for the external push and for friction, or are they different?

Is the amount that the block speeds up the same throughout the block's motion?

It was these types of activity-prompted clarifications that helped to generate the significant SMD found in the "Sliding blocks" sub-section in activity 3 of cycle 4.

---

[6] Research has shown that number-based activities tend to promote the rote application of mathematical procedures, so understand that these calculational activities are not meant to

Overall, although the purpose of this analysis was to explain the above-average percentage of time spent on SMD in "Sliding blocks", there is evidence throughout cycles 3-5 that supports my argument that diagramming, drawing, and graphing exercises scaffolded student analyses in such a way that their discussions were longer and richer than the discussions in activity sub-sections that did not make use of these activities. This is true even of group 2 (the "less-than conscientious" group), whose diagram-based conversations were generally richer than their conversations that did not revolve around diagramming or drawing. The main reason why more of these sub-sections don't show up in the top five list of sub-sections supporting SMD is that the graphing/diagramming/drawing exercises were often only a small portion of the sub-section, and so the large percentage of time that groups spend on SMD while diagramming/drawing/graphing disappear when they are averaged over the remainder of the sub-section.

Finally, although this is just a hypothesis, I would also suggest that students are aware that diagrams and drawings are the types of "answers" that are easiest to assess by the teacher. Inspecting and assessing a force diagram only takes 2 seconds, because it is largely a visual assessment; reading a word-only description of the same phenomenon, in contrast, may take minutes. Since students know that teachers can assess drawings easily,

---

fall into the same category as the "quantitative" diagramming, graphing, and drawing activities mentioned here.

they are perhaps more motivated to spend time working on and discussing activities that require drawing, diagramming, or graphing.

The group-related factor that supported SMD in the "Sliding blocks" sub-section in the Slowing Down activity was group 2's conversational excursion into the fine details of instantaneous speed. In this sub-section, group 2 spent over 5 minutes debating whether the slowing block loses its speed at a steady rate, or whether there is ever a time when the slowing block moves at constant speed. (This conversation has already been examined in the context of Result 1a, the result dealing with the fact that there were more instances of clarifying than expected in cycles 3-5.)

The interesting factor here is that, for no obvious reason, Roxanne chose to draw on her status as group leader to demand that the group help her work through her confusion about constant speed and slowing down. If nothing else, this speaks to the fact that an understanding-minded group leader can work wonders for sense-making discussion in his or her group. (Darla, in group 1, is also proof of this.) In other words, in group 2's case, the factor explaining the group's above-average engagement in SMD in "Sliding blocks" was only marginally relevant to the task at hand. Instead, the main question is why Roxanne -- a student who generally was far more concerned with drawing the group off-task than understanding the material, as shown below in Results 7 and 8 -- suddenly became motivated to construct a detailed understanding of motion and instantaneous speed. Regarding this question, however, I have no answers. It is certainly true that Roxanne experienced cognitive incongruity in this instance, which is one of Hatano's basic requirements for sense-making. But this was not uncommon for Roxanne; there were many points in the curriculum where Roxanne was confused. The issue of greater importance is that, in this one instance, Roxanne decided that working through her confusion was important (another requirement for sense-making) -- and I frankly have no idea why she decided that this particular confusion was something that needed to be worked through; her decision would have involved a judgment that her speed/motion confusion was somehow more interesting or more important than her other instances of confusion, and I have no insight to offer as to how she might have come to that conclusion.

Apropos of nothing, we now switch from the activity sub-sections with the highest percentages of time spent on SMD to the sub-sections with the lowest percentages of time spent on SMD.

"What really happens?" sub-sections from activity 1 in cycle 3 (6.5%), activity 2 in cycle 3 (4%), and activity 2 in cycle 5 (a.k.a "Is gravity caused by the earth's magnetism?"; 3%).

Because of the heavily constructivist stance of the curriculum developers, the developers erred on the side of caution when it came to letting students conduct experiments in order to construct their own understanding of the CIPS benchmark ideas. By this, I mean that developers allowed students to perform experiments that tested the validity of each and every benchmark idea (as well as the well-known misconceptions or alternative conceptions related to those idea) -- even in cases when the ideas of interest were relatively commonsense. As a general rule, it was also the intent that, in testing any benchmark idea, experiments should be kept relatively short and simple, and should give rise to results that are clear and unambiguous in their support or refutation of the idea (or related misconception).

From the students' perspective, unfortunately, this translated into a handful of experiments that had unambiguously obvious results (because they already knew what would happen), and were therefore unnecessary (because the students already had an intuitive understanding of the commonsense ideas explored in the experiment). In hindsight, then, it should probably not come as a surprise that these experiments -- as well-intentioned as they might have been -- did not support a significant amount of sense-making discussion.

Three of these "obvious and unnecessary" experiments were housed in the "What really happens?" sub-sections listed above (one per sub-section). Below is a more detailed description of these experiments.

- To see how forces combine, students weighed a number of individual objects, and then determined their combined weight (activity 1, cycle 3)

- As part of a single pulley system, students hung weights on opposite sides of the pulley in order to determine that the system moves toward the heaviest object(s) (activity 2, cycle 3)
- To test whether magnetism is a possible cause of gravity, students put a variety of objects near a magnet (to test their magnetic properties) and then dropped those same objects (to test their gravitational properties) (activity 2, cycle 5)

Following are some representative comments that illustrate the attitude that students took towards these experiments.

- "It's kinda stupid. Third grade stuff."
- "Why are we doing this? I mean, we know what the answer is."
- "(sarcastic) Gee, I wonder. Is a pencil magnetic?"
- "Everything's attracted to gravity! Duh!"

Clearly, the students did not view these experiments as useful or helpful in developing a new understanding of forces or gravity, respectively.

One thing that I want to make clear is that I'm not arguing that the ideas addressed by these experiments should have been skipped. For instance, it is very common for students to suggest that magnetism is a possible cause of gravity -- and so it is certainly valid for the class to explore whether this might be the case. It's just that an experiment (in hindsight) is probably not the best way to evaluate this idea. There's no reason why a brief (5-10 minute) guided discussion, with examples, couldn't have served the same purpose.

Recently, based on feedback from the curriculum pilot testing in California, Minnesota, and Michigan, the developers have come to realize that students don't need (or want) to construct ideas that everyone already "knows" to be true; a corollary to this is that students don't want to perform experiments whose results are easily predicted. Also, for the sake of interest and motivation, students seem to desire longer, more complicated experiments. As a result, the magnetism and hanging weight experiments have since been excised from the curriculum and replaced with developmentally more appropriate activities.

Finally, for the sake of completeness, I will briefly describe other problems with the above sub-sections beyond the "obvious and unnecessary" problem described above. In particular, in both the "combining weight" and "magnetism as a cause of gravity" experiments, there were activity-based distractions that drew the students' attention away from the core ideas of the experiment; due to these distractions, students ended up devoting their sense-making efforts to topics that were not directly related to the cycle benchmark ideas.

In the force/weight experiment, which was the experiment where students weighed objects both individually and as a group, students became very concerned with the accuracy of their measurements. The result of this is that students wasted their energies trying to determine the exact weight of the object(s), when -- from the developers' perspective -- students should have instead focused their sense-making efforts on the relevant benchmark idea: that forces combine additively. Of course, that this occurred was no fault of the students. Precision and uncertainly are both extremely important topics in physics -- they just didn't happen to be the focus of this particular exercise. The root of the problem is that the "hanging weight" experiment was one of the earliest experiments involving numerical values, and the developers had not anticipated that students would be concerned with the numerical precision of their measurements. A brief discussion about uncertainly and precision at the beginning of the activity may have significantly reduced the time that student spent weighing and re-weighing the weights to ensure that they were as accurate as possible.

In the magnetism/gravity experiment, the relevant distraction was the decision to have each group examine a compass at the beginning of the experiment. The sole purpose of the compass had been to demonstrate that the small magnet from the experiment behaved like the earth's magnetic field, but students in both groups became convinced that the compass was somehow an integral aspect of the experiment. Consequently, students in each group spent a considerable amount of time trying to figure out exactly how to use the compass (where to line up "north", etc.) Removing the compass from this experiment very likely would have solved some of the "distraction" problems in this gravity/magnetism sub-section.

"Prepare your wipe board!", from activity 1 in cycle 5 (2%). In Result 3, I carefully explained why groups failed to spend much time on sense-making discussions in "Prepare your wipe board!" sub-sections, so I will only summarize the reasons here.

Groups didn't spend much time on sense-making in these sub-sections because the revisiting and reconstruction of ideas over the course of a curriculum cycle did not occur as planned. Instead, students typically were satisfied with a particular idea long before the consensus activity arrived, and so the students did not feel the need to revisit, and therefore re-discuss, these ideas in the consensus "Prepare your wipe board!" sub-sections. Most of the time in these sections was dedicated to students dictating their already-finalized ideas to the person doing the writing on the presentation board.

Result 5: Sense-making Differences between Activities

Result 5: The percentage of time that groups dedicated to sense-making discussion was quite a bit higher in some activities than in others.

The purpose of isolating the CIPS activities with the highest and lowest overall small-group sense-making percentages is so developers might learn from the successes and failures of the CIPS pilot curriculum. Identifying the factors that contributed to these high/low percentages is a big step in ensuring that, in the future, these activities' successes might be duplicated, and their failures might be avoided.

The percentage of time dedicated to small-group SMD was highest (35% or higher) in the following activities:

Table 6-8. Activities with the highest percentage of time dedicated to sense-making discussion.

Activity Title	**Activity #**	**Cycle**
Slowing Down	3	4
No Friction?	4	4
More Exploration of What Causes Gravity	3	5

This percentage was lowest (10% or lower) in the following activities:

Table 6-9. Activities with the lowest percentage of time dedicated to sense-making discussion.

Activity Title	**Activity #**	**Cycle**
Lots of Pushes and Pulls!	3	3
Putting it All Together	4	3
What's a Little Drag?	5	4
Putting it All Together	6	4
What Causes Gravity?	1	5
Putting it All Together	5	5

In considering these high and low percentages, one thing to remember is that these percentages only reflect the times where students engaged in sense-making discussion in their groups. Excluded from these percentages are the many sub-sections where the class as a whole engaged in SMD. For example, the primary purpose of the Putting it All Together activities was for students to come to agreement in a whole-class discussion on the ideas that best explain the cycle phenomena; they did so in the "What ideas does your academy have?" sub-sections in those activities. Since the above percentages for the Putting it All Together activities only reflect the SMD from the "Prepare your wipe board!" sub-sections (which were the sub-sections just before "What ideas does your academy have?"), I cannot comment on the activities that best supported whole-class discussion. I can only comment on the activities that best supported small-group SMD, since that was the primary focus of this study.

Most of the factors that explain the high/low percentages in the above nine activities have already been addressed in other results in this chapter. Therefore, many of the factors will be re-described only briefly, without any supporting evidence.

<u>Slowing Down, activity #3 in cycle 4 (35%)</u>.

Activity summary: Students explored the effect of friction on motion. In teams, students slid blocks on their desks and explained why it slowed down. In so doing, they filled in energy diagrams and used other representations to describe the block's motion. They then reflected on the effect of friction on an object's motion.

There are both general and group-specific factors that explain the large sense-making percentage in Slowing Down, an activity that contained only one sub-section that supported small-group SMD ("Sliding Blocks").[7] The general factor -- a factor directly related to the curriculum -- is that "Sliding Blocks" required students to construct energy and force diagrams; comparing all activities, I found that activities with these diagrams provoked sense-making more readily than activities that did not require these diagrams (see Result 4). This support for sense-making was due to the conceptual scaffolding provided by the CIPS energy diagrams and the demand for precision required by the semi-quantitative (but not necessarily numerical) force diagrams and speed arrows. The group-specific factor for Slowing Down is that Roxanne atypically led her group (group 2) through an in-depth clarifying discussion of speed. The powerful effect of group leadership in this instance shows how extended sense-making discussions are readily supported when group leaders choose to make conceptual understanding a focus of their group (see Result 7).

<u>No Friction, activity #4 in cycle 4 (44%)</u>.

Activity summary: Students engaged in a class discussion about which sports might be possible without friction, and then they watched a Magic School Bus cartoon that showed what frictionless baseball would be like. Groups then considered reasons why they would or would not want to live in a frictionless world.

---

[7] The "Making sense" sub-section for this activity was skipped by both groups, and so "Making

This activity is one of the few activities where students actually enjoyed their small-group sense-making discussion (see Result 4). This alone is enough to mark this activity as extremely successful in comparison with the other activities in this study.

The reason for the interest and enthusiasm in these discussions is that, overall, the activity was a nice example of an intrinsically motivating activity, meaning that it had personal relevance, was novel, and gave the students some creative control over the activity.

<u>More Exploration of What Causes Gravity, activity #3 in cycle 5 (39%)</u>.

Activity summary: Students saw two videos. The first video recreated the Cavendish experiment, the experiment that shows how boxes of sand are attracted to bottles of water -- thus helping to prove that all objects have a gravitational attraction for one another. After the video, the teacher guided students through an interaction analysis (using the interaction tool) of the water and sand. The second video showed hammers dropping simultaneously on the moon and on earth. Students then reflected on the cause of gravity.

Three factors explain the high sense-making percentage in this activity: exemplary peer tutoring and interaction (in group 2), prompting in the curriculum materials, and cognitive incongruity.

The peer tutoring and interaction in group 2 was fueled by Arthur's interest in and knowledge of the phenomenon of gravity, as well as by Sabrina

---

sense" did not contribute to the overall sense-making percentage for this activity.

and Roxanne's cognitive incongruity related to the Cavendish video. After the video, Sabrina and Roxanne admitted that they were confused, so Arthur stepped in to help:

R I didn't know big things are attracted to small things.

A I get that, but I don't understand...it's weird.

S I didn't understand.

R Neither did I. How could water be attracted to sand?

A It's not that...it's just that...it's saying that if you have something sitting here, like something that's suspended by a string...(suspends pencil in his hand)... say this pencil was held by a string…if you put big old weights [near the ends of the hanging pencil], for some reason it [rotates towards the weights].

S I don't even understand the video.

R Big things are attracted to small things, I guess.

A It's basically saying somehow there's a force between...there's a little force around every object -- and the bigger the object, the bigger the force around it. Except for certain things, like...check this out. We've discovered these certain kind of planets that are this big (holds hands out -- beachball size). They have tremendous gravity. And the gravity is so strong that your hands...(on either side of imagined planet)...are just (moves hands to show that they are drawn together to hit the "planet")...like stuck to it. It's like you're wearing metal gloves and you have a mega-powerful electromagnet [pulling your hands to the small planet]. That's what it's like.

S And they're <u>planets</u>?

A Little tiny planets. They can be anywhere from the size of the moon, to the size of a baseball.

S And they're in outer space, or on earth?

A Outer space.

S Oh.

R I have a question for you, Arthur.

A Yes?

R How come...does Jupiter have gravity?

A (nods) Not much, but yes. I'm not sure.

R Well, if it's bigger...wouldn't it have more gravity, then?

A I'm not sure if that's the one...(Arthur turns to me, the reasearcher)...does Jupiter have more...have we discovered whether or not Jupiter has more gravity than earth? I don't remember. All the planets have gravity. I know that.

R Mr. Bohn? Does Jupiter have gravity?

T It certainly does.

A Is it more or less?

T More than the earth.

A (to R) You were right. (Jupiter has more gravity than earth)

A I want to know what causes the force around each object.

S Air pressure, I don't know. This is hard to understand.

A Yeah, I know.

R I get it. (laughs)

S I don't.

A (to R) What causes the force around each object? Do you get it?

R The size! (to A) Wait, what?

A Where does the object get its force. It's the...

S Air pressure?

A Does the moon have air on it?

S Why doesn't it?

A (to S) Does it?

S (to A) Yeah.

R Oh! The moon doesn't have lots of gravity because it's so small!

A No it doesn't [have air].

S Oh, is that why we need oxygen tanks?

A (to R) Yeah, but moon has gravity.

R But not a lot.

A I know. But where does it get its gravity from? How does it have it?

R Size.

S Friction! (A, R laugh) Ok, never mind. (S laughs)

Roxanne and Sabrina's initial cognitive incongruity, as seen in the excerpt, played a big role in initiating the extended SMD in this activity. More importantly, Arthur was willing to step up and explain in detail what he knew about gravity to Sabrina and Roxanne, and was furthermore willing to guide the group's discussion as to what the ultimate cause of gravity might be. Arthur's exemplary peer tutoring in this case was unusual, as he (as explained in Results 7 and 8, below) tended to be somewhat impolite in his dealings with his fellow group members. The factor that helped Arthur overcome his typical impoliteness, it seems, is that he was very knowledgeable and interested in the phenomenon of gravity, as indicated by his knowledge of neutron stars (small "planets" with lots of gravity) and his desire to learn about the cause of gravity ("I want to know about the force around each object").

The curriculum prompts that supported the SMD in this activity were found in the sub-section that compared earth gravity to moon gravity. In that sub-section, spaces were provided for students to record the prediction for each individual group member (which hammer will hit the ground first -- earth hammer or moon hammer?). The prompts took this form:

> ____________thinks: ________will drop first because
> ____
> ________________________________ .

The final sense-making prompt for the section was the following:

> After our discussion, our group thinks: ________will drop first because
> ____
> ____________________________.

The transcripts show that group 1 carefully followed these prompts, with the result that the prompts scaffolded the group's sense-making discussion in that the group was guided to (a) consider and evaluate the predictions from each group member as to which hammer will fall first (earth, moon), and then

(b) come to agreement on which prediction was most likely correct, and why. Because it was off-task, group 2 engaged in very little SMD in this section, and so the obvious conclusion is that these prompts are only valuable when groups feel obligated to follow and complete the written curriculum.

The other case of cognitive incongruity that had a significant impact on the SMD in this activity has already been described in Result 4. The incongruity stemmed from the fact that, in considering the experiment where two pencils are placed side by side (in "Making sense"), Darla recognized that gravity affects all objects, and so she simply could not understand why the pencils' gravitational forces did not move the pencils closer together. Darla's group leadership (discussed in Results 7 and 8) allowed her to guide her group through an extended sense-making discussion (over two days!), in which Darla attempted to clear up this confusion.

Lots of Pushes and Pulls!, activity #3 in cycle 3 (7%).

Activity summary: Students filled out force diagrams for multiple forces acting on an object in one dimension (representing tugs-of-war). Afterwards, groups reflected on the effect that multiple forces have on the motion of an object.

Force diagrams normally prompted nice sense-making discussions in CIPS, but they did not do so in this case because the teacher implied that the force diagrams should be filled out individually, and so there was very little discussion in this activity (see Result 1b). This is the primary factor behind the activity's low sense-making percentage (7%).

What's a Little Drag?, activity #5 in cycle 4 (10%).

Activity summary: Students watched a drag racing video, and then set up experiments to test how far cars travel with and without attached sails. They then reflected on the effect of drag on an object's motion and energy.

The factors that contributed to this activity's low sense-making percentage are: the results of the car/drag experiment were too easily predicted (see Results 2 and 4), the experiment itself took far too long to set up and run (and so the time dedicated to the experimental procedure far outweighed the time needed for sense-making), and the "Making sense" sub-section was skipped by both groups (see Result 1b).

What Causes Gravity?, activity #1 in cycle 5 (9%).

Activity summary: Groups debated the possible causes of gravity, and then the teacher led a class discussion on the same topic.

Although the whole-class discussions in this activity were exceptional, the small-group discussions were not. Group 1's discussion was unexceptional because the group members spent most of its time reading their ideas aloud (an activity not considered to be sense-making discussion). The sense-making that did occur was extremely brief because Darla's idea was almost immediately recognized as being the "best"; it was recognized as such because Darla was the only group member to offer reasonable supporting evidence for her idea. Group 2's discussion was unexceptional because the group chose this particular activity as a time to be off-task, which meant, in this instance, that group 2 didn't choose to take the activity and its associated small-group discussion very seriously. These factors are discussed in more detail above, in Result 4.

Putting it All Together, final activities in cycles 3-5 (8%, 10%, 1%).

Activity summary: Groups recorded their final ideas on presentation boards, and then the teacher led a class discussion as to which ideas best explained the cycle phenomena.

Groups didn't spend much time on small-group sense-making in these activities because the revisiting and reconstruction of ideas over the course a curriculum cycle did not occur as planned. Instead, students typically were satisfied with a particular idea long before the consensus activity arrived, and so the groups did not feel the need to revisit, and therefore re-discuss, these ideas in the consensus activities. Most of the group effort in these activities was dedicated to students dictating their already-finalized ideas to the person doing the writing on the presentation board. These factors are explained in more detail in Result 3.

Result 6: Sense-making Differences between Curriculum Cycles

Result 6: The percentage of time that groups dedicated to sense-making discussion was lowest in cycle 3.

During this study, participants worked through three cycles of CIPS Unit 2 (Interactions and Motion):

- Combining Pushes and Pulls (cycle 3)
- Resistive Interactions (cycle 4)
- Non-contact Interactions -- Gravity (cycle 5)

It was found that a smaller percentage of group activity was dedicated to sense-making discussion in cycle 3 (12%) compared to cycles 4 (22%) and 5 (17%). Why is this so?

Some of the characteristics of these cycles are identical. For example, each cycle followed a similar structure: elicitation activity, then development activities, and then a consensus activity. Also, each cycle was written by the same curriculum developer. Therefore, it is unlikely that the structure or style of these cycles had any relevance to the difference in sense-making percentages.

There were two main differences between these cycles, however. The first differences is obvious: the cycles focused on different topics in physical science. Cycle 3 pertained to combining forces, cycle 4 explored friction and drag, and cycle 5 was aimed at verifying and refuting the various possible causes of gravity.

The second difference is not so obvious. Data that points to this "hidden" difference is that group 1's percentage of time dedicated to SMD actually reached its peak during cycle 3 (30%), while group 2's percentage hit its all-time low during cycle 3 (6%). It's only because the values from groups 1 and 2 were averaged together that the overall percentage for cycle 3 was so low. Therefore, the main reason for the differences between cycles is not ground in the differences in topics, but is instead tied to group 2's extremely small cycle 3 sense-making percentage -- a value that is fully three times less than group 2's second-lowest cycle percentage (18%, for cycle 5). Accordingly, the real question that gets at the differences between cycles is the following: Why did group 2 spend so little time on sense-making discussion in cycle 3?

The primary difference between cycles 3-5 for group 2 is that, for half of the activities in cycle 3, group 2 had an additional member: Jasper. Jasper was unique in that, in the short time that he participated in this study, he didn't engage in a single instance of verbal sense-making. The vast majority of his time was spent off-task with Roxanne, and during the few times that he was on-task he tended to be loud, distracting, and verbally abusive. In terms of the roles described in Chapter 2, Jasper was a clear example of an active noncontributor: the person who engages in large amounts of off-task behavior and also challenges and ridicules the other group members.

The main problem with the group's SMD in cycle 3 is not so much that Jasper was in the group, however. The real problem is that both Jasper <u>and</u> Roxanne were in the group together. In cycles 4 and 5, when Roxanne was the only group member primarily interested in off-task behavior, group 2 was able to successfully engage in significant SMD under the following circumstances: (a) when Roxanne chose (for whatever reasons) to be on-task, and perhaps even lead the group through SMD, and (b) when Arthur and Sabrina chose to ignore Roxanne's off-task overtures. In contrast, when both Jasper and Roxanne were in the group -- that is, when group 2 contained <u>two</u> group members who were primarily interested in off-task behavior -- group 2 became an off-task juggernaut that could never find a way to engage in significant SMD. Basically, Jasper and Roxanne's joint off-task antics were nearly always successful in either (a) drawing one or both of the remaining members (Arthur, Sabrina) off-task, or (b) forcing Arthur and Sabrina to work silently, in isolation from the rest of the group. These effects of Roxanne and Jasper's off-task behavior contributed to the fact that group 1 didn't engage in significant SMD when both Jasper and Roxanne were in the group, and so its percentage of time engaged in SMD during this time (most of cycle 3) hit an all-time low.

The following transcript from activity 2 in cycle 3 (Lots of Pulls!) illustrates the extremely negative influence that the Jasper-Roxanne pair had on group 2's SMD. Immediately prior to this excerpt, the team had skipped the prediction section of the activity (because it hadn't been paying attention) and had done some pulley experiments with legos to see how combined forces (i.e., stacked legos on opposite sides of the pulley) affect the motion of the system. Sabrina had to leave earlier in the period, so Arthur was left alone with Roxanne and Jasper.

	(J tosses lego to A; J messes with the lego, and R messes with the pulley)
A	(puts legos on his side of system and lets system go; string moves toward A)
R,J	(talk off-task)
R	Teract is a woman. That's why she's so strong.
R,J	(talk off-task)
A	(puts system back on pulleys)
A (to R)	Roxanne.
R	What?
A	Push [the string].
R (to A)	Now what do we do? Are we done?
A	Actually, we were supposed to write everything we just did down.
R,J	(grumble and groan)
R	All right.
J	(talks off-task)
R	(reads) "What do pulls combine when they are in opposite directions? What is your team's reasoning?"

J (jokes) Opposites attract!

R (reads) "How do pulls combine when they are in opposite directions? What is your team's reasoning?"

A (to R) Wrong [question]. (Roxanne was looking in the wrong question on the worksheet because she hadn't been paying attention)

A (points out to R where they are) There.

R (turns page, reads) "Predict what will happen when you let go of the paper clip [attached to the string, and at the center of the pulley system], and write your reasoning."

A (to R) (points at a different question) Read that first.

R (reads) "One team member should hold the paperclip steady while the other team members attach Teract's legos to one pulley and Kinet's legos to the other..." Ok. (writes) "We predict that Teract..."
(note that they should have done their prediction <u>before</u> the experiment)

A (reads) "Teract's side will drop because it's [bigger], which means it's heavier."

R Yeah. (writes) "We predict that Teract's strengths will win...win because...because..." Because why? Because she has more strength?

A Because that side's heavier.

R Ok. (writes) "Because that side's heavier."

J (still goofing around with lego)

R (reads) "Holding the paperclip, set up the tug-of-war so that it's Stas and Kinet against Teract."

J (finally opens up his notebook 30 minutes into the class period)

R (reads) "Predict what will happen..."

R (to A) Hey, what did we predict? (looks at A's paper)

A It will stay the same because both sides have the same weight.

J What the truck? Where are you guys at?

R (writes) "Stay the same...

(J asks A about worksheet, A tells him which questions to answer)

A (to J) Basically, what we ended up with is that Teract's side will drop because it's heavier.

R (writes) "...stays the same...because the weight is even." (reads) "Now suppose Stas, Kinet, and Modulas pull together. Predict what will happen when you let go of the paperclip, and write your reasoning."

R,J (talk off-task; R drops lego and breaks it, then R, J continue their off-task discussion

J (to A) What did you put for [question number] three?

J (taps R) What did you put for three?

R (turns back page, reads) "Teract's side drops because it's heavier. "

J (speaks to a member of another group) Do you want to just bring your notebook over here so I can copy you?

This excerpt is typical of the chaotic, distracting activities and discussions that took place when both Jasper and Roxanne were in the group together. Here, Arthur was left to make sense of the activity alone, and Roxanne was willing to write down anything that Arthur told her to write.

Jasper, Roxanne's partner in off-task crime, was even willing to copy the answers from another group's notebook.

As you can see, having two group members that were primarily concerned with being off-task was disastrous for this group. This is why group 2 engaged in extremely little SMD during cycle 3, which in turn brought down the overall cycle 3 sense-making percentage. As soon as Jasper left the group -- who was the one group member who had absolutely no interest in CIPS -- group 2 was able to significantly increase its sense-making discussion.

Result 7: Sense-making Differences between Groups

<u>Result 7: The percentage of time dedicated to sense-making discussion was higher in group 1 than in group 2.</u>

Relevant explanations and assertions:

- Group 1 felt obligated to stay on-task, complete the activities as intended, and fill in all of the appropriate blanks. In general, group 2 did not feel these same obligations.
- Group 1 was more concerned than group 2 about understanding the underlying concepts of the CIPS curriculum.
- The leaders of group 1 (Darla and Lacey) encouraged the group to cooperate, complete the activities as intended, and evaluate and understand the concepts underlying the CIPS curriculum. The leader of group 2 (Roxanne) was almost always looking for an

excuse to go off-task, and frequently succeeded in drawing other group members off-task.

- In group 2, one of the main contributors to sense-making discussion (Sabrina) maintained an expectation that verbal and written explanations should be simple and brief. The main contributors in group 1 did not share this same expectation.
- In group 2, there was friction between the two main contributors of sense-making statements. There was no friction between the main contributors in group 1.

When participating in groupwork, groups 1 and 2 engaged in a number of different activities. These activities included writing answers, setting up of experimental apparatus, recording data, and engaging in various types of discussion: negotiations of meaning, actions, status, materials, and also off-task discussion. Comparing the percentage of time that these two groups engaged in SMD provides a measure as to how much time each group spent (relative to all of these other activities) on discussion of meaning -- discussions which, ideally, were meant to help students understand and develop ideas about the physical world.

The result from Chapter 5 was that group 1's percentage of time dedicated to SMD (26%) was 63% higher than group 2's percentage (16%). There are a number of group characteristics that explain this difference in group SMD, each of which is described in detail below.

Off-task behavior. An important difference between groups 1 and 2 was in the frequency and length of their off-task behavior. In cycles 3-5, the members of group 1 were nearly always on-task, meaning that they spent all of their time doing the things that they were supposed to be doing without letting anything distract them from the task at hand. They generally only stopped working when they had completed an activity sub-section and were forced to wait before moving on to the next sub-section (until the rest of the class finished the current section, for instance). Group 2, on the other hand, engaged in off-task behavior many times over the course of the study. In any given activity, there was a fair possibility that group 2 (or at least some of its members) would be off-task.[8] The overall result of this off-task behavior was that group 1's percentage of time engaged in SMD was higher than group 2's percentage because (a) group 1 spent more time than group 2 engaged in sense-making discussion, and (b) group 1 often spent less time than group 2 in each activity sub-section, since group 2's off-task activity would often draw out the time that they spent on each section. More plainly, since group 1 generally engaged in more sense-making discussion in less time, its sense-making percentage was higher than group 2's percentage.

Quantitatively speaking, group 2 was off-task as an entire group for one minute or more on 14 separate occasions. Adding these off-task times

[8] Group 2's off-task behavior was often visible enough that the teacher would stop by to remind the group to stay on-task and work at a faster pace. Also, on two occasions, group 2 was so noisily off-task that they attracted the attention of the entire class.

together, group 2 was off-task during cycles 3-5 for a grand total of 21 minutes. The longest time that group 2 went off-task as a whole was for 3 minutes, in activity 3 from cycle 3 (Lots of Pushes and Pulls!). The highest overall percentage that group 2 was off-task was also in Lots of Pushes and Pulls!, in which the group was off-task 63% of the time (12 out of 19 minutes). During cycles 3-5, group 1 was never off-task as an entire group for over a minute.

Not included in these figures are the instances (many instances, in the case of group 2) where "only" a few members of the group were off-task. When Jasper was still in group 1, for example, there were many times when both Jasper and Roxanne were jointly off-task -- playing with equipment, discussing the goings-on in other classes, revisiting their weekend plans, and so forth. Once Jasper left the group, there were still numerous occasions where pairs of students in group 2 (Roxanne-Sabrina, Roxanne-Arthur, or Arthur-Sabrina) would go off-task for long periods of time, although the most likely off-task pair was Roxanne-Sabrina. In terms of going off-task alone, Roxanne was by far the most likely to go off-task all by herself, during which time she might sing, attempt to draw the others off-task (sometimes successfully), or talk to members of another group. In group 1, the most likely pair to go off-task was Lacey-Grace, although these two only went off-task a few times. Lacey was the member of group 1 most likely to go off-task by herself, the purpose of which was usually to chat with a member of another group.

In this study, when most of a group was off-task (2 out of 3 members, e.g.), there was very little possibility of the last remaining group member engaging in SMD. The sole on-task group member might record his or her results or answers, or engage in silent reflection, but the discussion of scientific concepts with the remaining group members was largely out of the question. Group members could sometimes engage in SMD when only one group member was off-task, but even one group member's off-task behavior proved to be a significant hindrance to SMD at times -- especially when that behavior was successful in drawing other group members off-task.

Completing the activities as intended, and filling in the appropriate blanks.

One might hypothesize that a group's off-task behavior is but one symptom of a bigger problem: that the group has a lackadaisical attitude toward its science class. In addition to off-task behavior, such an attitude would probably manifest itself in a number of different ways:

- a consistent failure to read or follow instructions
- not turning in homework
- the tendency to leave some worksheet questions unanswered

Keeping these things in mind, consider the following facts about group 2's interaction with cycles 3-5 of the CIPS curriculum materials. In these cycles, Group 2…

- asked questions of the teacher when the answers could be found in the written instructions for the activity
- invented experimental predictions for the Lots of Pulls! activity after-the-fact (i.e., after conducting the experiment) because they weren't paying enough attention to realize that they were supposed to think about and record their predictions before running the experiment
- sometimes waited for teacher instructions to move on to the next activity sub-section, even when the teacher had already instructed groups to complete that sub-section as part of the day's activities
- frequently failed to turn in their homework
- left an average of 1-2 worksheet questions blank each activity

These facts support the hypothesis that group 2 was a lackadaisical group. Taken together, the characteristics above paint a picture of a group that is not terribly motivated to do well in the CIPS classroom, and would therefore not be likely to engage in extraordinary amounts of sense-making discussion -- as was already verified, above, by group 2's tendency to slip into off-task behavior.

<u>A concern for answers instead of understanding</u>. In a CIPS classroom, it would be fair to assume that a certain percentage of students begins the

school year with an expectation that CIPS, rather than being about the careful examination and construction of scientific concepts, is all about getting answers down on paper. One reason for this expectation is that, unless they live in a particularly progressive school district, students enter the CIPS classroom having already experienced many years of traditional instruction. And since the traditional classroom tends to value vocabulary words, facts, and correct answers over conceptual understanding, students with a background heavy in "traditional" experience would have no reason to assume that CIPS would be any different. Another reason why students might focus on answers rather than understanding in CIPS is a general lack of interest in science or school (or both). These students would likely be doing what they could to just get through the course with the minimum required work -- where "minimum work", in this case, would translate into putting answers on the CIPS worksheets so that the teacher has something to grade.

Given the real possibility that some students will primarily be concerned with getting answers instead of developing a new scientific understanding, one of the challenges of CIPS is to engage students to the extent that they are genuinely concerned about the examination, debate, and development of their ideas. Students who successfully shift from an answer- to an understanding-based focus would, one hopes, occasionally be able to set aside their need for answer-scribbling in favor of the sharing, exploring, and evaluating of ideas that is so crucial to CIPS.

As related to this study, student perceptions of CIPS as being answer- or understanding- based have a direct effect on group sense-making. Groups that are focused on answers would be expected to settle for early versions of an idea without giving much thought to its validity or usefulness, copy down ideas without seriously evaluating them, and perhaps even look to authority (either within the group or without) for correct scientific ideas. As a result, groups that focus on answers would be expected to engage in less SMD than groups that focus on genuine conceptual understanding, a focus that would be reflected in their percentages of time engaged in SMD.

Looking at the two groups, one of the factors that explains the difference in the two groups' sense-making percentages is precisely the issue raised here: understanding vs. answer-getting. One group (group 1) was concerned with both understanding and answer-getting, in that the group was definitely concerned about filling in their ideas and explanations on their worksheets, but they typically did so only after the group was satisfied with the validity of these ideas and explanations. Because of its lackadaisical nature, the other group (group 2) was less focused than group 1 on filling out the worksheets in the first place -- but when the group was actually focused on filling them out, there were many instances where the group appeared to be primarily concerned with putting down answers instead of developing new understandings.

Group 1's concern for understanding is well-documented in the section on group leadership, below, and so only the answer-based focus for group 2 is documented here.

Evidence for group 2's focus on answers instead understanding is when group 2…

- decided not to record its own answers for the cycle 3 consensus activity, but instead presented a wipe board left over from a previous class period
- wrote that "friction prevents the pencils from moving" to the question "Why don't two pencils on a table attract each other with their gravitational force?" (activity 3, cycle 5), even though the originator of the question (Arthur) admitted that his explanation was the first thing that came to mind, and a total guess
- recorded a "yes" answer to "Did you observe the same motion in the dropped paper ball as the simulation shows?" in the Gravity and Motion activity only because a "no" answer required additional writing -- even though the group's observation didn't match the results of the simulator

Other supporting evidence lies in the fact that group 2 asked sixteen answer-related questions during cycles 3-5 (e.g., "What do I put here?", "What's the answer?"), whereas group 1 only asked six answer-related questions during the same period of time. That is, while group 1 more commonly asked sense-making questions, group 2 more commonly asked questions of the "What answer did you get?" variety. Asking someone to recall their written answers was specifically excluded from the general category of "sense-making" because, more often than not, the point of such questions is to obtain someone's answer (to a worksheet question, e.g.) without understanding it.

Clearly, based on the above evidence, group 2 tended to be less focused on understanding, and more focused on answer-getting, than group 1. However, it would not be proper to argue that group 2 was solely focused on getting and writing down answers in their CIPS notebooks. Indeed, there were many instances of group 2 becoming engaged in lengthy sense-making discussions about forces, energies, and gravity. But it is quite clear that there were times when group 2 focused on answers over understanding, the result of which is that the group's SMD suffered during these times. Along with all of the other factors described in this section, group 2's occasional answer-based focus helps to explain why group 2's percentage for SMD was lower than group 1's percentage.

<u>Group leadership</u>. Treating a group as a single undivided unit, as is sometimes done in this dissertation, is certainly something that should not be done lightly. Obviously, groups are composed of individual students, and individual students have their own ideas, experiences, and attitudes. Therefore, when a claim is made that a particular group felt obligated to

complete activities as intended, or was answer-focused, or wasn't interested in a particular science topic, there is necessarily a degree of generalization and averaging that goes into assigning these attributes to that group. In group 1, for example, Grace was much more answer-focused than the rest of her group, yet students in group 1 still mainly focused on the understanding of ideas. Likewise, Sabrina was the one student who typically tried to keep group 2 on-task, although there were many times when the entire group went off-task -- the result of which is that group 2 is characterized as a group that didn't feel obligated to stay on-task and complete the activities as intended.

One of the reasons that groups can be treated as self-consistent units with their own characteristics is that all group members do not have equal status within the group. Some students have more control over the group's actions and discussions than the others, for a variety of reasons. These students, who are referred to as ***group leaders*** , can (and do) use their influence to guide the way that the group interacts, thinks, and learns. Accordingly, with the current research project in mind, one would expect the leader to have an effect on his or her group's SMD.

In this study, the two groups were under the influence of different types of leaders, with the effect that one group was guided to engage in much more SMD than the other. Specifically, Darla and Lacey used their joint leadership of group 1 to guide its members into cooperating, completing the activities as intended, and evaluating and understanding the concepts underlying the CIPS curriculum. On the other end of the scale, Roxanne used her leadership of group 2 to frequently distract the group of its sense-making efforts by drawing its members off-task.

The first way that Darla and Lacey positively influenced their group was to keep alive a spirit of cooperation and group-mindedness. Through Darla and Lacey's words and actions, students in group 1 were periodically reminded that the group was doing something together: considering an idea, evaluating hypothesis, making a prediction, or whatever else the curriculum required. One way that Darla and Lacey achieved this goal was to occasionally begin their sentences with the word "we":

"So, do we think that…" (Darla)

"So, we're saying…" (Darla)

"What do we think…" (Lacey)

Using the word "we" in this way helped cement the idea that the members of the group should be thinking and learning together, that anyone in the group might contribute to the discussion, and that the final ideas are meant to represent the group consensus. The spirit of group-mindedness with similarly maintained when Darla followed up Lacey's prediction in the Can You Lend Me a Hand? (activity 1, cycle 3) with "Everybody agree with that?" Here, Darla again used her group leadership to emphasize that the entire group should be in agreement on the ideas and predictions produced by the group.

Lacey also made comments that promoted cooperation, although she tended to be more direct and proactive in making sure that the group worked and moved through the activity together. For instance, when Porter inevitably began working silently and individually (as outlined in Result 8, below), Lacey would sometimes chastise him by saying "It's teamwork! Teamwork!", meaning that he should not be thinking and writing in isolation from the other members of the group. Lacey took a similar attitude toward Grace when she finished a section in the Lots of Pushes and Pulls! Activity (activity 3, cycle 3) before the rest of the group had caught up, and Grace began asking questions about the next section. Lacey's response to Grace was "I'm not there yet! We have to wait for our teammates." In What's a Little Friction?, in the next cycle, Lacey made a similar comment when Grace again tried to move ahead of the group: "Wait! I still have to do this part." The implication of both comments is that the entire group should finish a sub-section before anyone in the group should be allowed to move forward.

Another way that Lacey's group leadership helped to achieve cooperation and group-mindedness was the way that Lacey read many of the worksheet questions aloud to the group. After she cheerfully read a question, Lacey typically looked up and paused to show that she was genuinely expecting an answer. These reading/looking signals for group conversation nearly always prompted someone in the group to suggest an answer to the worksheet question; other comments would often follow, and so Lacey's method of reading questions aloud commonly ended up eliciting extended sense-making discussions. An interesting contrast is that, when members of group 2 read the questions aloud, they often did so incompletely (trailing off at the end of the question, e.g.) or under their breath (mumbling, etc.), which made it obvious that the reader was mainly reading aloud for his or her own sake, and didn't necessarily expect an answer from the other members of the group.

While the spirit of cooperation and group-mindedness laid a solid foundation for the members of group 1 to work and think together, there were other more direct ways that group leadership supported SMD in this group. One example is the way that Darla would not settle for an idea that was unclear, problematic, or confusing, and would get the entire group involved in testing and modifying the troublesome idea. This can be seen in Darla's constant requests and demands for elaboration, clarification, and explanation from her fellow group members. Examples of this are shown in the transcript excerpts, below. Implicit in these excerpts is that Darla accepted that her group members were useful resources in helping her develop a better understanding of the concepts in the CIPS curriculum.

Example 1: Considering the question of how pulls combine

D (reads) ok..."how do pulls combine when they are in opposite directions?" Do they combine? Because they're going against each other.

G No.

L Well...

G Well if you're like pushing the same thing...I don't know. I'm not saying...

D I don't think they combine, because they're...

G They're all in the opposite...

D ...pulling in opposite directions. Can they actually combine them? I mean, you put more on this side (one side of pulley) and more on that side (other side of pulley)...

G It's on the...they're pulling the same row (?) so...

P It's almost like negative weight. This has positive weight (one side of pulley), this has negative weight (other side) -- so you add them together… I don't know. (shakes head)

D So, do we think they can be combined or not?

P, L, G No.

Example 2: Considering the length of force arrows due to friction

D (looking at her force diagram) So, we drew a longer arrow [on the block with rougher sandpaper].

L Yeah.

D (reads) "Explain the length of the push arrow that you drew." We...that's only because the sandpaper is rougher, so it's not...

L [The rougher sandpaper] has a stronger resistance.

D	Yeah, but would [the force arrow with the rougher sandpaper] be shorter or longer? Because if there's more resistance...(looks at diagram again)
L	I think that it would be shorter if it makes more resistance.
D	No, I think it's longer. (pause) Yeah, longer. Because there's more resistance on the rough...
G (to D)	So we put that on the bottom?
D	(nods) It's only because there's more resistance.
L	So we drew it longer because there's more resistance.
D	Yeah, between the sandpaper -- the rougher sandpaper -- and the wipe board.

Other examples of Darla's push for comprehension are seen in the group's discussion of why pencils don't gravitationally attract each other (shown in Result 4, above) and the group's filling out of an energy diagram for a slowing block (also shown in Result 4). Clearly, Darla's constant questioning and guidance prompted reflection and reconsideration of the ideas in the minds of her fellow group members, the result of which the group likely engaged in much more SMD than if Darla had not been in the group.

One last way that Darla's and Lacey's group leadership supported group SMD was in how they pushed the group to strictly follow the activity instructions and work through the proper order of questions and directives within the curriculum. For example, in Lots of Pulls! (activity 2, cycle 3), Darla stopped the group from jumping ahead into the lego/pulley experiment by pointing out that "Ok, we have to predict [before we can go on]." She made a similar comment in Gravity and Motion (activity 4, cycle 5), when she again directed the group's attention to a question in the written materials: "…we're supposed to answer this question before we [set up the experiment]. So let's just answer that." Final examples of this are when the group haphazardly started messing around with magnets and objects in Exploring What Causes Gravity (activity 2, cycle 5), and Lacey guided the group back onto the right track by saying "We don't know what we're doing. We have to read [the instructions]." -- and also later when, in the same activity, Lacey pointed out in the instructions that "We have to discuss [this question] as a team." Without Darla and Lacey's guidance to follow the curriculum as written, the group likely would have skipped over many questions that were intended to prompt reflection and group discussion.

In terms of the leadership styles outlined in Chapter 2, Darla and Lacey's style of leadership was closest to the *inclusive* style. Typically, leaders using the inclusive style bring up ideas, ask group members for their opinions, and carefully consider their input. Clearly, Darla and Lacey did these things -- but labeling their leadership style as "inclusive" seems to fall short of the true extent to which Darla and Lacey actively promoted cooperation, the completion of activities, and the tenacious evaluation and understanding of the concepts underlying the CIPS curriculum

Now we turn to group 2, which was under the leadership of Roxanne. Roxanne was an extremely talkative and energetic student who would almost always look for an excuse to go off-task. For example, Roxanne might spend a significant amount of time each day doing one or more of the following: singing songs, discussing events that happened during lunchtime or over the past weekend, talking to other groups, complaining about spelling or numbering errors on the worksheet, making funny noises, speaking in goofy accents, and so forth.

Roxanne was the group's leader in the sense that, like Darla and Lacey in group 1, her words and actions had a large influence on the behavior of the rest of the group. For example, Roxanne was frequently able to draw other group members away from their work by engaging them in off-task activity.[9] On the positive side, though, Roxanne twice used her leadership within the group to draw the other group members into extended, in-depth discussions of phenomena or concepts -- but the use of Roxanne's leadership to promote group sense-making discussion was extremely rare.

Specific examples of the undesirable effects of Roxanne's leadership on the group are as follows:

- Roxanne convinced the group to present a wipe board during the cycle 3 consensus activity (activity 4) that had been written by a group in a previous class period

[9] For example, Roxanne once suggested that the group should purposely spill a glass of water so that the group could waste time cleaning it up.

- Roxanne urged her fellow group members to skip the "Making sense" section in the Slowing Down activity (activity 3, cycle 4), and they did so
- Roxanne recorded that "god" and "magic" were the causes of gravity during the What Causes Gravity? elicitation (activity 1, cycle 5), which completely deflated the seriousness of the activity and kept the group off-task for the entire period

Unsurprisingly, under Roxanne's leadership, group 2 didn't engage in nearly as much sense-making discussion as it could have.

Comparing the two groups, many of group 1's group-minded and comprehension-oriented characteristics are notably absent in group 2. Members of group 2 rarely, if ever, used the word "we" as they spoke to one another. As already mentioned, members of group 2 tended to read aloud to themselves, rather than read aloud to the rest of the group. It has also already been established that group 2 was less-than-extraordinary in its willingness to follow directions; it was not uncommon for the group to skip sections or particular questions, for example. (Arthur or Sabrina would occasionally point out something in the curriculum that needed to be done, but these types of comments were fairly rare.) Also, it was not terribly uncommon for the members of group 2 to be on different pages of the same activity. Lacey tried to keep the members of group 1 all on the same page, but there was nobody to fulfill the same role in group 2. Finally, in group 2, there was no real equivalent of Darla's constant requests and demands for elaboration, clarification, and explanation. Typically, Arthur was the group member who would be relatively confident in his answers, and Sabrina would be confused, but still interested in clearing up her confusion. But because Roxanne rarely exercised her leadership to promote sense-making, Sabrina was left with Arthur to clear up her confusion -- and, as discussed below, Arthur and Sabrina really didn't get along very well.

As you might imagine, the many differences in the leadership styles of groups 1 and 2 contributed significantly to the differences in the two groups' percentage of time dedicated to sense-making discussion. If group members are being drawn off-task, then they're not engaging in SMD. And if a group's members are on separate pages in the same activity, tend to skip over sense-making questions, and just plain don't get along, you can imagine that the group would be less likely to engage in SMD than a group that does not share those same characteristics.

<u>Why can't we all just get along?</u> In chapter 5, it was recognized that interpersonal skills would likely play a key role in the small-group discussion that was meant to be the primary support for student reflection and learning in CIPS. In fact, interpersonal skills did turn out to be one of the factors that had a significant impact on the SMD in this study.

A notable difference between group 1 and group 2 is that the members of group 1 tended to be much more polite, socially aware, and non-judgmental than the members of group 2. As you might imagine, the effect of these differences on sense-making discussion was profound. Group 1's discussions flowed and evolved smoothly, while the conversations in group 2 were sometimes not allowed to run their full course because they would degenerate into snappy, bickering exchanges in which someone (usually Arthur) would cop an attitude or make a personal attack.

As exceptional examples of politeness and social awareness, Lacey and Darla (the leaders of group 1) would sometimes preface their contributions to group discussions with phrases like:

"Do you want to hear what I wrote?" (activity 1, cycle 4)

"Do you want to hear what I put down?" (activity 2, cycle 4)

"Wanna hear [my idea]?" (activity 1, cycle 5)

These statements helped to create an atmosphere of politeness and respect that helped the members of group 1 contribute freely to the sense-making discussions in cycles 3-5. Moreover, when the members of group 1 took turns presenting their ideas to the group, the other group members were always very non-judgmental about these ideas. By this, I mean that students never made comments to the effect that these ideas were dumb or silly, or that the comments indicated that the person might be stupid. The members of group 1 certainly felt free to criticize and challenge a contributed idea (Darla would often do so as soon as the idea was presented the group), but these criticisms or challenges were never meant to be personal attacks on the individual who had constructed the idea. The following, taken from the elicitation activity on the cause of gravity, is an example of this sort of polite, non-judgmental exchange:

G	[The cause of gravity is that] there's gravity on earth that lets things fall and not, like, float around.
D	Right...so you...basically you said that gravity pulled it to the ground.
G,L	Yeah.
L	I said that gravity, like, has something to do with the atmosphere, or something. Wanna hear? Mine is...I think gravity is...um, something in the earth's air that...the earth's air has magnets. Like little magnetism things in it that pulls things to the ground. (shrugs)
D	If they're little magnets, wouldn't they be in the earth to pull things down?
G	Maybe it's only the ground that makes things go down. (laughs)
L	Maybe there's something in the core...at the core. You know how no one can ever get past the...outer surface, you know? Maybe there's something in the core that pulls everything into...onto the ground.
D (to G)	What do you think?
G	I put that there's a force around the earth that makes things go down. Like, it keeps all the air inside, so it could fall...
D	The atmosphere?
G	Yeah, the atmosphere. It just sits staying there and just float around there.
L (to D)	Ok, what'd you put?

D I put that I thought gravity was caused by the atmosphere because...uh, once you go out of the earth's atmosphere, there's no gravity.

G Yeah, and makes things float around.

The sense-making discussion in group 2, in contrast, took on an entirely different character than the type of discussion shown above. The direct result of Roxanne's off-task leadership is that Roxanne typically was not available (by choice) for the sense-making discussions in group 2. Roxanne's attentions would be elsewhere, and so Sabrina would be forced to turn to Arthur for help in resolving any conceptual difficulties that she was experiencing.

In the day-to-day functioning of group 2, Arthur's general role was that of "idea man". Both Sabrina and Roxanne were aware that Arthur's ideas typically proved to be fruitful, and that Arthur had some interest in science -- which, in turn, meant that Arthur had knowledge of science that they did not possess. For example, Arthur once excitedly recounted for Sabrina and Roxanne an experiment that the saw on TV, in which a feather and baseball were dropped in a vacuum -- and they both fell at the same rate. On another occasion, Arthur tried to tell Sabrina what he knew about the black holes and neutron stars. Overall, because of his interest in and knowledge of science, Arthur attained a special "idea man" status within the group such that Sabrina and Roxanne (when she chose to participate) typically ran their ideas past

Arthur to see what he would think, and if he had comments of his own. The unfortunate reality, however, is that there were times when Arthur was neither polite nor respectful in his role.

It is very difficult to depict Arthur's bouts of impoliteness with transcript excerpts, as his rudeness was nearly always a matter of tone.[10] At various times, especially in his conversations with Sabrina, he could be condescending, belittling, dismissive, and even angry. However, to give some indication of what these discussions were like, I present the following transcript of Arthur and Sabrina's back-and forth exchange (activity 4, cycle 5) regarding an interaction tool (Figure 6-2) and follow-up questions, tonal comments included. The interaction in this case was between a dropped paper ball and the earth.

[10] It was not always a matter of tone, however. Arthur was not above saying things like "your idea makes absolutely no sense at all" or (half-jokingly) "Roxanne is stupid and a moron".

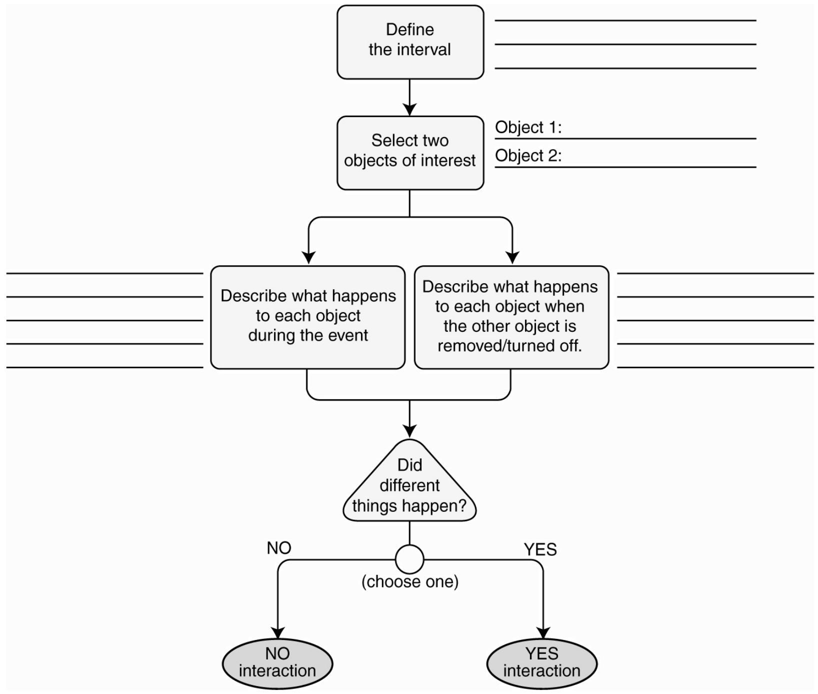

Figure 6-2: CIPS Interaction Tool

S Ok. (reads) "Define the interval." Ok, what do we put...(reading)

S (writing) "[The time interval] starts when the..." (stops writing, turns to A) ...paper ball drops?

A (condescending) It starts from when the paper ball leaves your hand and...until it hits the ground.

S (to A) [Object 1 is] the ball...and [object 2 is] the table? Or earth?

A (dismissive) Ground, table, whatever. First one's paper.

S (indignant) I know that.

S (reads) "Describe what happens to each object during the event." The paper ball...

A (irritated) The paper drops and hits the table, the ground, or whatever or whatever.

S What does the ground do? It doesn't do anything! (Arthur disinterestedly turns his head and says: hmmm?)

A If the ground was removed, the paper would just...(makes dropping motion with hands)...keep falling and falling and falling and falling. And then if the paper was removed, nothing would hit the ground.

S (annoyed) Roxanne, do you have all this?

(R copies S's interaction decision tool. S reads her answers to her.)

S (reads) "Did different things happen?"

S (to A) (raising her voice to get A's attention) Hey, did anything happen? No?

A (curt) Yes. (S grimaces at his tone)

S (reads next set of instructions) "Each person in your team should drop the paper ball a couple of times. Everyone should watch it fall."

S (still reading) "What kind of motion does the paper ball have when you drop it? Speeding up, slowing down, constant speed?"

R I think it speeds up. Because, you know how...Apollo 13 (the movie...), how they let go -- and then it started to burn up on [re-entry]? So wouldn't that mean that it speeds up?

S (to A) Well, what do you think?

A (disinterested) I have absolutely no idea.

S (annoyed) No idea at all. No guesses at all.

A Nope.

In this transcript, notice that Roxanne's attentions were elsewhere, and so she was unavailable for discussion; she paid so little attention, in fact, that she was forced to copy the answers from Sabrina's interaction tool. The bulk of the sense-making discussion, with Roxanne effectively absent, therefore involved only Sabrina and Arthur. During this discussion, there were many opportunities for Arthur to guide Sabrina towards a better understanding of the interaction between a dropped ball and the earth, such as when Sabrina questioned Arthur about which objects were interacting (the paper ball, earth) or when the interaction interval started and began (from the time when the ball drops to when it hits the ground). But Arthur's tone and attitude prevented Sabrina from taking advantage of these opportunities to better understand the troublesome gravitational interaction.

The above exchange between Sabrina and Arthur was one of the worst as far as interpersonal skills are concerned. Therefore, you should not assume that Arthur and Sabrina's discussions were always this impolite and unhelpful. There definitely were instances when Arthur guided Sabrina towards a better understanding of a troublesome idea or phenomenon (and vice versa). Yet, in most instances, the discussions between Sabrina and Arthur were ground in a minor undercurrent of friction and rudeness, and so their discussions never quite reached a level of politeness that might be called "pleasant". Compounded with the fact that most of the sense-making discussion in group 2 was left to Sabrina and Arthur, the sense-making discussion in this group was definitely not as extensive or in-depth as it might have been. In other words, the oft-seen friction between Arthur and Sabrina was one of the factors that contributed to group 2's sense-making percentage being lower that group 1's percentage.

<u>Expectations for the length and depth of explanations</u>. From a scientist's perspective, some of the basic features of scientific explanations are clarity, precision, and completeness. One might even argue that the purpose of sense-making discussion is directly tied to these features. In SMD, one hopes that student comments would help to construct an explanation that, over the course of the discussion, would become ever more clear, precise, and complete. But what of simplicity, and brevity? The philosopher Occam became famous for his rule (Occam's Rule) that the simplest of competing theories should be preferred to the more complex, and -- to some extent -- this rule even carries over to the domain of science in that brief, elegant descriptions and explanations are sometimes preferred to descriptions and explanations that are more convoluted in nature. However, there are some instances when simple, brief explanations can obscure, ignore, and even misrepresent the true underlying explanations for a phenomenon or experimental result. Therefore, people need to be careful in striving for simplicity and brevity when constructing scientific explanations -- especially when that person is a student who still needs to develop an understanding of the finer details of these explanations.

In group 2, the tension between simplicity, accuracy, and completeness was in full force, and had a definite impact on the group's sense-making discussion. In this group, Arthur was generally the proponent for longer, more in-depth explanations, while Sabrina (and sometimes Roxanne) argued for shorter, simpler explanations. Consider, for example, group 2's discussion from activity 1 in cycle 4 (Will It Slow Down?), at which point students were asked to consider whether a pushed skateboard (on earth), bicycle (on earth), shopping cart (on earth), and wrench (in space) would keep moving at a constant speed or eventually slow down. Specifically, students were asked to predict the outcomes for the skateboard, bicycle, cart, and wrench, and then provide reasoning for their predictions.

An excerpt from this discussion:

A A skateboard will always slow down because of the friction on the wheels. A bicycle is the exact same thing. And exact same with the shopping cart. Because there's no gravity, or wind, or anything to block [the wrench]…[the wrench] can go on forever.

A (writing his answer) "Because of gravity, the friction between the wheels...and ground..."

R (reads) "You are on a skateboard in the middle of an empty, flat parking lot."

S (reads) "Will you move along without slowing down?" No.

A (reads his answer) "Because of gravity, the friction between the wheels and the ground causes the skateboard to slow down."

S That's too complicated. "No, because you can't move forever."

forever."

R No, because of gravity.

S (writes) "You...can't move forever."

In this example, it is interesting that Sabrina's response to Arthur's in-depth explanation was not to critique it, or evaluate whether it was actually correct -- but instead was to discard it outright in favor of a much simpler (and much less accurate) explanation. If Sabrina had been confused by Arthur's answer, for instance, this would have been the perfect time to engage Arthur in SMD about the validity or accuracy of his answer. Instead, Sabrina decided to completely forego discussion in favor of recording a simpler (and much more brief) explanation.[11]

Another example of Sabrina's preference for simple and brief explanations is in group 2's answering of Idea #2 from cycle 3 ("How do combined pushes and pulls affect motion?") on their presentation board during the cycle 3 consensus activity. During the related discussion, below, Roxanne was in charge of recording the idea on the presentation board, and Arthur and Sabrina argued over the exact form of the answer.

S Roxanne, [write this]. "In a game of tug-of-war, when you combine people…"

A ...people/pulls...

[11] An interesting factor in this conversation is that, based on her comments later in the cycle, there is some evidence that Sabrina believed that "all things slow down" is actually an innate property of objects. I.e., It is possible that "all things slow down" was a causal explanation in Sabrina's mind, while a causal explanation for Arthur was something completely different (friction on wheels, etc.).

S ...the stronger side wins.

A ...the combined force of either side will cause motion. (repeats) "The combined force or strength of either side will cause motion."

S (to A) I don't even understand what you're saying. I was understanding what I was saying.

R (reads) "...combine people plus..."

S What? That doesn't make any sense.

A (telling R what to write) "In a game of tug-of-war, when you combine people or pulls, the combined strength of either side will cause the [system]…"

S Just say: "When the pulls...pulls the combined strength...the stronger side wins."

A You have to say something about motion. "The stronger side wins" doesn't mean anything.

S Well...it's too complicated what you're saying.

R The stronger side...strongest side has a more effective motion...

A The stronger side causes the...box in the middle [of the rope] to move.

R (to A) I thought you said we have to mention motion.

A That is... ...causes the box to speed up.

S "In a game of tug-of-war, when you combine people/pulls, the stronger side causes the box to move."

A To speed up...to one side. To speed up towards it. You can't put "move" because you have to mention the fact that it's speeding up. That's mentioning motion.

In the transcript excerpt, Arthur's concern for completeness and accuracy "won out" over Sabrina's desire for uncomplicated explanations in the sense that Arthur got his more complicated explanation on the presentation board, but what did not occur as a result of this exchange was for Sabrina to explore, challenge, or really understand his answer.

In part, the whole purpose of SMD is for students to discuss things that they don't understand, and so it is frustrating to find a situation where a student was reticent to understand an explanation just because it is "too long" or "too complicated" (in her view). This exchange had been an opportunity for Arthur to guide Sabrina towards a better understanding of net force and motion, but this did not occur -- primarily because Sabrina's response in this instance was to say "what you're saying is too complicated" instead of to say "I don't understand that -- what do you mean?" The lack of SMD in this instance means that, most likely, Sabrina left the discussion still at a basic level of understanding ("the stronger side wins", or "the rope moves"), a level of understanding that was less precise and less sophisticated than Arthur's. Nobody would argue that Sabrina's answer was wrong (it was technically correct, on some level), but Sabrina's level of explanation was certainly much lower than it should have been.

To summarize, in explaining the result that group 2's percentage of time engaged in sense-making discussion was quite a bit lower that group 1's percentage, Sabrina's willingness to settle for brief, simple explanations was one of the factors that contributed to this result. Many opportunities arose where Sabrina could have asked Arthur to elaborate on one of his "complicated" ideas -- which likely would have provoked an extensive sense-making discussion -- but Sabrina typically chose to ignore those explanations in favor of keeping her own simpler, shorter explanations.

Result 8: Sense-making Differences between Individual Students

Result 8: In both groups, the amount and type of sense-making statements per group member varied quite a bit.

Ideally, in an inquiry-based classroom with a large emphasis on groupwork, each student in class participates to the fullest extent possible in the many small-group discussions that are prompted by the structure of the curriculum. Group members who do not participate frequently in small-group discussions (or only participate in particular ways, such as by only asking questions) are not able to fully benefit from the group's collective wisdom. Group members who do not share their ideas cannot have these ideas evaluated and critiqued so that they might become more precise and more complete, and group members who do not comment on others' ideas are depriving their fellow group members of guidance that could prompt those group members to change their ideas for the better.[12]

[12] Another reason why students should not withhold their guidance from others is that the very act of guiding (pointing out perceived flaws, etc.) provides the "guider" with yet another opportunity to reflect on and reorganize his or her own ideas, which in turn might help the guiding person's ideas evolve even further. In more colloquial terms, sometimes the best way to learn is to teach!

Since it is desirable for all CIPS students to participate equally in small-group discussions, one would hope that the distribution of sense-making instances would be roughly identical for each student. That is, the desired result would be that students engage equally in the six components of sense-making discussion: predicting, connecting, testing knowledge compatibility, and so forth. The actual result of this study, however, is that the eight students ended up with very different distributions of sense-making instances. Why this was so is described in detail below.

<u>Group 1: Darla</u>. Darla has already been described as one of the leaders of group 1: a leader who helped her fellow group members cooperate, complete the activities as intended, and evaluate and understand the concepts underlying the CIPS curriculum. In directing the group 1's sense-making discussion, she wouldn't settle for an idea that was unclear, problematic, or confusing, and would instead involve the entire group in testing and modifying the troublesome idea. This can be seen in Darla's constant demands for elaboration, clarification, and explanation -- demands that were placed both on herself and on her fellow group members. Another aspect of her leadership is that, although not emphasized up to this point, Darla was more than willing to engage in peer tutoring, at which times she took on the role of student-teacher for her fellow group members. An example of this is from the "Making sense" sub-section in Lots of Pushes and Pulls! (activity 3, cycle 3), in which students answered the following question: "How does the total push or pull on an

object affect the object's motion?" The following is a transcript excerpt from this sub-section:

L Ok, let's see. (reading what she's written) "It affects an object by giving it power to change its speed."

L (to D) What did you put?

D I put...I think you're supposed to say how does it change speed, like speeding up or slowing down. I put: "The total pull or push makes objects speed up or..." -- like if they're the same like when Stas and Kinet and and Teract were pulling -- "...they are...they maintain a zero speed. Because if we have a pull that's the same, then it's a different thing. Unless, all of a sudden, somebody has, you know, a sort of energy and all of a sudden it goes (makes sound, moves hands like pulling something).

L (to D) Is it "affects", or "changes"?

D "Changes".

L (writes) "Changes...changes its motion...by..."

L (reads) I put: "Changes motion by when people -- in parentheses, 'strengths' -- pushes or pulls on something." (shrugs)

D Ok, but you probably want to say something about speeding up, or making it constant...er, maintaining zero speed. Because, on here, remember, "objects can:" -- and then they give you this list. They want to know like how it affects it. Like, does it speed up...

L Ok.

As you can see, Darla's gentle reminders that the question was directed at the object's motion are excellent examples of peer tutoring in action.

Quantitatively, Darla's peer tutoring and frequent demand for group sense-making are both represented by the fact that Darla had more instances of providing an underlying explanation, describing/defining/connecting, and testing knowledge compatibility than the other members of her group.

Of course, *why* Darla took on the role of "director of sense-making" is a different matter entirely. One very likely explanation for her behavior is related to the concept of student goals, a concept that was introduced in Chapter 2. Applying the goal framework to Darla's behavior, it would be reasonable to argue that Darla frequently possessed a *learning* goal, which translated into her frequent attempts to gain a deeper understanding of the concept or task at hand. One of the factors that contributed to her possessing a learning goal (as opposed to possessing a performance goal, for example) is that Darla was one of the few students to appear to have a genuine interest in science. Her interest can be seen in the way that she sometimes went beyond the script of an activity to perform some additional experiments of her own. For example, in activity 1 from cycle 3 (Can You Lend Me a Hand?), Darla extended the original activity (weighing groups of objects to see how forces combine) to include the weighing of her own pencil. The only other student who went beyond the boundaries of the written curriculum to perform additional experiments was Arthur, another student who I argue (below) had a general interest in science.

I leave as a final description of Darla the teacher's comment that Darla was "the best student in the class -- very sharp, and very hardworking".

Grade-wise, Darla's performance relative to the other students can be seen in the table below.

Table 6-10. Cycle 3-5 power drive (test) scores and quarter grades.

		Power Drive Scores				
		Cycle 3 (28 pts)	Cycle 4 (15 pts)	Cycle 5 (20 pts)	Quarter 2 Grade	Quarter 3 Grade
Group 1	Darla	26	9	16	A	A-
	Grace	7	6	0 (actual score)	C+	C
	Lacey	25	7	9	B	B-
	Porter	?	--	--	A	--
Group 2	Arthur	24	8	16	A-	A-
	Roxanne	23	10	9	B-	B-
	Jasper	0 (not turned in)	2	5	F	D
	Sabrina	14	0 (not turned in)	6	C+	C+

Note. Cycle 3 contributed to students' 2nd quarter grades. Cycles 4 and 5 contributed to students' 3rd quarter grades. Jasper left the study after the first week of cycle 3. Porter left partway through cycle 4; his notebook was thrown away immediately after he left, so Porter's power drive score is unavailable for cycle 3.

Group 1: Lacey. Along with Darla, Lacey was one of the leaders of group 1. Lacey was an energetic, talkative student who was fairly conscientious about staying on-task and learning the concepts underlying the CIPS curriculum materials. Lacey's verbal sensemaking was supported by Darla in that Darla was always available for SMD, and also because Darla would draw Lacey's attention back to the written materials when Lacey occasionally went off-task to discuss topics with students outside the group. Overall, it appeared that Lacey found an equal amount of enjoyment in both the academic and social aspects of school.

The major difference between Lacey and Darla's distributions of sense-making instances lies in the difference in definitions, descriptions, and connections: 25 instances for Darla, and 15 for Lacey. The explanation for this difference is that, although the leadership of group 2 was jointly shared by Lacey and Darla, Darla usually served as the ultimate authority and provider of guidance when it came to sense-making in this group. Even though Lacey and Darla jointly sponsored the sense-making in group 2, Lacey often turned to Darla for final comments, elaborations, and explanations (as illustrated in the Darla-Lacey example of peer tutoring, above).

Because of her concern with learning and understanding, it would be fair to say that, like Darla, Lacey often possessed a learning goal. However, I don't feel that I can make any strong claims about Lacey having a general interest in science, since there was not as much evidence for this being true of Lacey as there was for Darla. For instance, Lacey was never seen to go beyond the written materials to extend an activity. However, it is true that she rarely appeared bored by any of the activities; also, she readily admitted that she was a fan of the Magic School Bus television science show.

Group 1: Grace. Based solely on Grace's distribution of sense-making instances (Table 5-9), one might assume that Grace had engaged in as much sense-making discussion as either Lacey and Darla. Unfortunately, this was not the case. Grace's distribution is the one distribution that is misleading in this regard.

Not reflected in Grace's distribution is the fact that it was either Darla or Lacey (or both) who typically drove the group's sense-making discussions. They generally did so with challenging conceptual questions and/or conceptually elaborate statements. Grace's sense-making statements, while nearly equal in number to Darla's and Lacey's, were most often brief side comments on Lacey and Darla's back-and-forth conversation. Even more illustrative of Grace's role within group 1 were her questions, which often were not sense-making requests at all, but were instead requests for answers. Overall, the breakdown of roles within group 1 were that Lacey and Darla played the roles of initiator and sustainer of the group's SMD, and Grace's role was typically limited to that of commenter and "answer seeker".

The following transcript illustrates Lacey, Darla, and Grace's typical roles in sense-making discussion. Here, the group attempts to answer a question in the "What do you think?" sub-section[13] of activity 3 of cycle 3 (Lots of Pushes and Pulls!). The question: "Suppose you have a combination of

[13] This sub-section was not included in the analysis of sense-making discussion because the questions in this section were supposed to have been filled out individually by the students.

pulls on an object, such as the tug-of-war shown below. How do you think the combination of pulls will affect the object's motion?"

D Well, see...because these guys were stronger than Teract by herself, so the combination of pulls made her (Teract) lurch forward.

G (to D) And fall into the mud.

P I don't know.

D (to P) It depends which object they're talking about. I mean, that's an object with ropes and all.

G So, what do we put?

L Wait wait! The combination of the pulls will affect the object's motion by...the people with the more pull will cause the other person's motion to go forward instead of being able to pull back.

D Right.

G You pull instead of pushing.

D Like what Porter said. Which object are they talking about? Are they talking about this little thing in the middle of the rope, or the rope, or the people themselves?

G The tug-of-war. The whole tug-of-war. The people and the rope.

L Yeah. The whole fact about the tug-of-war.

D Ok.

G Because "you have a combination of pulls on an object, such as in the tug-of-war shown below."

D Ok, well...the combination of pulls will...that's kinda hard to explain what it will do. I mean, it will affect...the motion. Because it will make one of them either move forward...and the other one move backwards.

G And fall.

D Yeah.

G So, what do we put?

In this example, Darla initiated the sense-making discussion with a prediction about the effect of the forces on Teract, and then sustained the discussion with a question about the object of interest (the thing in the middle of the rope, the rope, or the people?) and a description of how the forces might affect the motion of the system ("I mean, it will affect the motion…"). Lacey also contributed significantly to the SMD with an elaboration on Darla's prediction ("the people with the more pull will cause…"). Grace's comments were limited to clarifications of the physical phenomena ("And fall into the mud"), brief conceptual elaborations ("You pull instead of pushing"), re-readings of the worksheet question ("you have a combination of pulls on an object, such as in the tug-of-war shown below "), or requests for an answer ("So, what do we put?"), none of which were key in initiating or sustaining the group's SMD. Lacey and Darla's roles of initiators and sustainers of the group's SMD, and Grace's roles of commenter and "answer seeker", as evidenced in this transcript, were maintained over the entire length of the study.

Framing Grace's behavior in terms of goals, Grace's primary goal during groupwork was to obtain answers from her fellow group members by either looking at their papers or asking them about their thoughts and opinions. She usually did her best to fill in all of the blank spaces in each activity's worksheet, although it is clear that she often did so without reflecting on the answers' significance or underlying concepts. Her low scores on the power drives (including a zero on the cycle 5 power drive, as seen in Table 6-10) demonstrate that she developed little understanding of the cycle 3-5 benchmark ideas. Additionally, when asked to put her ideas to paper, it was extremely common for Grace to look at either Darla's or Lacey's papers for guidance on what to write.[14]

In terms of the roles outlined in Chapter 2, Grace was probably closest to the role of "passive noncontributor", who is described as the person who rarely participates in group activities and often copies their work from others. Yet, this description is inaccurate in the sense that Grace almost always participated in the activities (indirectly, at least) -- it's just that, after she did so, she resorted to the copying of answers that is typical of the passive noncontributor.

Compared to the others, the explanation for Grace's behavior is more uncertain. But, after my constant observation of Grace for a month-long period of small-group activity, I am convinced that the main factor behind Grace's behavior is that the level and speed at which Grace could contribute

[14] I really cannot emphasize this point enough. On the videotapes, Grace was seen to constantly check Lacey and Darla's papers in order to copy what they wrote or compare her answers to theirs.

to her group's discussions were so far below the intellectual level/speed of her fellow group members that she could not participate significantly in the sense-making in her group. Based on her actions, statements, and extremely poor test performance, I believe that Grace's sense-making contributions were typically limited to side comments and questions because, intellectually speaking, Lacey and Darla's explanations, hypotheses, and reasonings were so far beyond Grace's own apparent level of thinking that making side comments was the only way that Grace could participate. Grace's direct attempts at sense-making (explanations, connections, etc.) were nearly always fragmented and difficult to follow -- so much so, in fact, that Lacey and Darla (as group-minded as they were) began to ignore her contributions by the end of the last cycle. More often than not, her contributions were ignored because, to Lacey and Darla (based on their reactions), they either weren't relevant or didn't make sense. Toward the end of the study, in fact, because Grace rarely had anything significant to contribute in terms of group sense-making, it even reached the point where Lacey and Darla would completely ignore her statements, and sometimes even speak as if Grace were not present.

I don't mean to say that Grace had a negative experience in group 1. As I already pointed out, Lacey and Darla were generally very supportive of "whole-group" sense-making in their group, and this extremely tolerant attitude lasted for almost the entire duration of the study. In their day to day interactions, Grace, Darla, and Lacey were really quite friendly. It wasn't until the very end of the study that Grace became slightly marginalized in the group, meaning that Grace's contributions (and constant requests for answers) grew less and less bearable to the main sense-making contributors of group 1: Darla and Lacey. What really frightens me, in all honesty, is how Grace might have fared in any other group besides group 1. I find it difficult to image how Grace would fit into a group that lacked the tolerance, respect, and cooperation that was typical of group 1's interactions.

Group 1: Porter. As seen in Table 5-9, Porter was second only to Jasper in terms of his extremely small number of instances of verbal sense-making (14 total for Porter). However, Porter remained in the study for almost two full curriculum cycles, while Jasper participated in the study for only a few days. The main reasons for Porter's lack of verbal-sensemaking were that (a) he preferred to work alone, and (b) when he did desire to contribute, his past history of working alone, coupled with his lack of communication and interpersonal skills, interfered with his attempts at engaging in sense-making discussion with his group.

Porter was a quiet, thoughtful student who was only in his first or second year of public school. (Porter had been homeschooled for most of his life.) In class, he would often work ahead and complete the activity's written work long before the rest of his group. Porter was extremely confident in his ability to answer the CIPS worksheet questions, and would often point out that the curriculum was too easy. When unsure of an answer, Porter would sometimes look at Darla's paper or ask Darla her opinion on the correct answer. Overwhelmingly, though, Porter would sit quietly and write without making any effort to engage the others in conversation, the effect of which was that Porter made himself a stranger in his own group.

Lacey and Darla sometimes attempted to draw Porter into thinking and talking with the rest of the group, but -- since these attempts were usually unsuccessful -- Lacey and Darla eventually learned to leave him alone. Once, noting that Porter was engaged in his typical silent writing, Lacey chanted "It's teamwork! Teamwork!" in an effort to get Porter to speak with the group. Porter's response? "That's why I go to homeschool -- so I don't have to work as a team" (activity 1, cycle 4). There were numerous instances where Lacey or Darla asked Porter a question, only to have him completely ignore the attempt at contact. Other times, Porter would respond with "figure it out yourself". Plainly, Porter's interpersonal skills interfered with his ability to engage in SMD with the rest of his group.

The few times that Porter tried to explain something to the other members of group 1 were met with unfriendly responses, since Darla and Lacey already harbored some resentment from Porter's past history of working in solitude. This problem was compounded by the fact that Porter lacked the communication skills necessary to clarify his novel ideas for the rest of his group.

Consider, for example, the following transcript from the Lots of Pulls! activity (activity 2, cycle 3), in which the group tried to determine whether pulls can be combined. Lacey begins the discussion by noting that P is again working by himself.

L [Porter writes an answer] without discussing it with the group. It's a group [activity]!

D (to P) [Are you writing] a note to your girlfriend? (P shakes his head) Then what are you writing?

P My guesses.

D What are you guessing?

P It's like a negative variable added to a positive variable... it's almost like negative weight.

D Yeah, right.

P This has positive weight (one side of pulley), this has negative weight (other side) -- so you add them together,…

D Whatever.

P I don't know.

D Oh well, we don't understand it.

L (to P) Could you talk a little louder?

P This weight (one side of pulley) added to this weight (other side of pulley).

D Right.

P So this pull (one side) is less...

G It'll go down.

D Oh, I get it.

P But its pull is less, it's just…

L I still don't really get it.

P Like if there was nothing [on one side of the pulley] then it would go down faster. But there's something there, so it goes slower.

D That's kinda hard to explain, though.

L No it's not. If I knew what it was, I could explain it. But I can't. (P makes a face)

D (to P) What? I'm trying to understand here. I don't.

In this excerpt, Lacey's opening comment about Porter again working alone, together with Darla's "whatever" response, illustrate Lacey and Darla's initially unforgiving and cold attitude toward Porter. His past history of working alone clearly worked against him in this case. Furthermore, Porter could not find a way to put his novel explanation (two forces pulling in opposite directions are like a negative weight combining with a positive weight) in terms that Lacey and Darla could understand.[15] This example is typical of Porter's inability to communicate his novel explanations in simpler terms.

<u>Group 2: Roxanne</u>. As the main provider of off-task leadership in group 2, Roxanne's sense-making statements were less frequent and, in general, were conceptually less elaborate than the sense-making statements made by

---

[15] Another possible explanation is that, because of his past behavior, Lacey and Darla didn't particularly want to understand Porter's novel idea.

either Arthur or Sabrina. The reason that Roxanne had as many sense-making instances as she did is that there were a few isolated incidents where Roxanne engaged in extraordinary amounts of sense-making.

As explained above, Roxanne was an extremely talkative and energetic student who would almost always look for an excuse to go off-task. For example, Roxanne might spend a significant amount of time each day doing one or more of the following: singing songs, discussing events that happened during lunchtime or over the weekend, talking to other groups, complaining about spelling or numbering errors on the worksheet, making funny noises, speaking in goofy accents, and so forth. Moreover, Roxanne was frequently able to draw other group members away from their work by engaging them in off-task activity. In a very few rare instances, however, Roxanne used her leadership within the group to draw the other group members into extended, in-depth discussions of phenomena or concepts. This is the primary reason why, in group 2, the number of sense-making instances for Roxanne is second only to Arthur's sense-making instances.

Two of these Roxanne-intensive discussions were 2 of the 3 unexpected clarifying discussions from Result 1a. These unexpected discussions focused on (a) whether a slowing object (due to friction) move at constant speed and then slow down, or just slows down, and (b) whether air drag and air friction are the same thing. Had Roxanne somehow been convinced to participate more fully in the CIPS curriculum, I have no doubt that

Roxanne would have led her group to record-breaking levels of sense-making discussion.

Group 2: Sabrina. Sabrina made more requests for sense-making than any other student in this study. This is because Sabrina was the member of group 2 who felt most obligated to complete each activity's worksheets, keep the group on-task, and understand the activities' underlying concepts.[16] Sabrina and Arthur were the two main contributors to group 2's SMD; when she became confused, Sabrina would turn to Arthur for clarification and guidance. During these instances, Sabrina typically would pose a question to Arthur, consider his response, and then resume her questioning.

Numerous examples of this are found in the following transcript (from Slowing Down, activity 3 in cycle 3).

S Ok. (reads) "How does a push or pull affect an object's motion?"

A If we're correct, then a push or pull can...start the...start and stop the motion.

R (reads) "How does a push or pull affect an object?" An object's motion.

A (to himself) How does a push or pull affect an object's motion?

A Basically, there's a push or pull...is...can start and stop the object...

S Start and stop?

[16] Thanks to Roxanne, there were also a fair number of days where Sabrina was willing to go off-task for extended periods of time.

A Yeah. It can start and stop an object's motion. Right this second I can't think of a way to say that, because my brain's not functioning properly.

S So, it would be...object's motion...

A (writing) "An object's motion can be affected by pushes and pulls by...because..."

S Because?

A I know it, I know it. I just don't know how to say it.

.
.
.

S (reads) "A push...affects an object's motion?"

A An object's motion is affected by pushes...

S Wouldn't it depend on how hard?

A Yeah, but...is affected by pushes and pulls because...

S Without them it won't move.

R It's a reaction! Pushes and pulls react the dang thing to move.

A It can cause...it can cause the object to either start or stop. It can cause the motion

S Well, you've got to push it -- and then it will go.

A Yeah. A push can cause an object's motion.

R Well, what if you don't push it and it doesn't move?

A That's not being affected by pushes or pulls.

S So a push or pull will affect it by...starting it...

A By either starting it, stopping its motion, or reversing it.

S Wouldn't it be something about increasing and decreasing speed?

A If you want to put those in.

R But if, like, the thing you're gonna push weighs a lot. And you try to push it but it won't go?

S Why wouldn't it?

This transcript typifies the question-and-answer exchanges that resulted in Sabrina making many requests for sense-making, and Arthur engaging in many actual instances of sense-making.

Group 2: Arthur. Like Darla, Arthur was a student with a fairly high interest in science. He was interested to the point where he would sometimes extend the written activities to perform his own experiments, and would sometimes even try to engage me (the researcher) in on-topic conversation when no other group member would respond to his questions. Interestingly, he also felt somewhat obligated to help me collect useful data, a fact that was made obvious when Arthur encouraged his group (at least twice) to stay on-task and "do science" for the sake of my research study.

Generally, Arthur was not the initiator or director of sense-making in his group, so much as a source of knowledge ("idea man"). Arthur was usually quite happy to work on and fill out the CIPS worksheets on his own; it typically took Sabrina's questions and confusion to draw him into sense-making discussion. These discussions between Arthur and Sabrina, however, were only somewhat productive in that (as described above, in Result 7) there tended to be friction between Sabrina and Arthur. Arthur was actually extremely lucky in the sense that Roxanne and Sabrina didn't hold his impolite behavior against him from one day to the next. Unlike Lacey and Darla's intolerance of Porter's solitary behavior, Sabrina and Roxanne were never afraid to re-approach Arthur on a day following one of his bad moods. Finally, like Sabrina, there were times when Arthur was willing to go off-task for extended periods of time.

Arthur's role as an idea man with an interest in science is reflected in his distribution of sense-making instances, in that Arthur had more instances of verbal sense-making than any other member of his group. As you would expect, many of these instances of verbal sense-making were direct responses to questions from Sabrina (more likely) or Roxanne (less likely). Some instances of sense-making, however, were generating during Arthur's excited lectures to Roxanne and Sabrina on space, neutron stars, gravity, and frictionless freefall.

<u>Group 2: Jasper</u>. Jasper was unique in that, in the short time that he participated in this study, he didn't engage in a single instance of verbal sense-making. The vast majority of his time was spent off-task with Roxanne, and during the few times that he was on-task he tended to be loud, distracting, and verbally abusive. In terms of the roles described in Chapter 2, Jasper was a clear example of an active noncontributor: the person who engages in large

amounts of off-task behavior and also challenges and ridicules the other group members.

Jasper appeared to have little interest in understanding the principles of science or completing each day's activity -- often to the extent that he had no idea where the group was in the day's activity. A common tactic was to let his groupmates do all of the work for an activity, and then copy the worksheet answers from a member of the group (or from a neighboring group, even). After the first week of cycle 3, Jasper was relocated to a separate group of "troublesome" students.

Summary

The purpose of this chapter was to pinpoint the factors that contributed to the eight quantitative differences in sense-making discussion (between groups, sub-sections, etc.) that were identified in Chapter 5. To summarize the factors contributing to the SMD in this study, I first present a result-by-result summary of factors, and then reorganize the analysis by factor type: group factors, personal factors, task factors, and contextual factors.

Analysis Summary, by Result

Result 1a: There were more instances of clarifying a phenomenon or experimental result than expected.

The primary factor contributing to this result is that groups held unexpected clarifying discussions about the topics of gravity, instantaneous speed, friction, and air drag. These discussions were necessary because they explored conceptual details that were not specifically addressed by the CIPS curriculum.

Result 1b: There were fewer instances of defining/describing/connecting than expected.

The four sub-section types that were meant to support this component of SMD in CIPS were: "What does your team think?", "Making sense", "Now what does your team think?", and "Prepare your wipe board!".

There were fewer instances than expected of this component of SMD because (a) "Making sense" and "Now what does your team think?" sub-sections were frequently skipped by at least one group, (b) group 2 students worked on one "Making sense" sub-section individually, instead of as a group, and (c) "Prepare your wipe board!" sub-sections contained less sense-making than expected.

The "Making sense" and "Now what does your team think?" sub-sections were frequently skipped because of a lack of class time, a lack of student interest or awareness in completing these sub-sections, and a lack of formal teacher support for these sub-sections.

Group 2 worked on one "Making sense" sub-section individually instead of as a group because, in general, both the teacher and students were somewhat unclear as to which sub-sections should be worked on in groups, and which sub-sections should be worked on individually.

The factors that explain why groups didn't spend as much time on sense-making as expected in "Prepare your wipe board!" sub-sections are provided in the summary of Result 3, below.

Result 2: More than half of student predictions were not supported by evidence.

Predictions were often not supported by evidence. This was the case because (a) some predictions were too simple, and so the student's reasoning was obvious enough that it didn't need to be stated, (b) students twice misinterpreted "assuming this to be true, what should happen?" predictions to be of the "predict what will happen" sort instead, (c) student were occasionally interrupted by their fellow group members in mid-prediction, and (d) student predictions were sometimes the result of guesswork, meaning that there were times when students had no clear reasoning behind their predictions.

Result 3: The percentage of time that groups dedicated to sense-making discussion was lower than expected in "Prepare your wipe board!" sub-sections.

Groups didn't spend as much time engaged in sense-making discussion as expected in "Prepare your wipe board!" sub-sections because the revisiting and reconstruction of ideas over the course a curriculum cycle did not occur as planned. Instead, students typically were satisfied with a particular idea long before the consensus activity arrived, and so the students did not feel the need to revisit, and therefore re-discuss, these ideas in the consensus "Prepare your wipe board!" sub-sections. Most of the time in these sections was dedicated to students dictating their already-finalized ideas to the person doing the writing on the presentation board.

Group 1 was satisfied with their ideas early in the idea development process because they thought that their ideas were basically correct, and no one in the group was in a position to provide the guidance necessary to clarify or improve these ideas.

Group 2 was satisfied with their ideas early in the idea development process because they didn't value the ongoing self-examination of their ideas.

Also, there were instances when a cycle idea didn't need redevelopment because the teacher hinted strongly at (and sometimes even explicitly told) students the correct cycle idea.

<u>Result 4: Groups spent a large percentage of their time engaged in SMD in certain individual sub-sections, and a very small percentage of their time engaged in SMD in others</u>.

Sub-sections in which groups spent a large percentage of their time engaged in SMD had one or more of the following characteristics:

- they were intrinsically motivating, meaning that they were personally relevant, novel, and gave students creative control over the activity
- they required students to construct force and/or energy diagrams

- they provoked cognitive incongruity in a group member who was both willing and able to lead the group through an extended sense-making discussion
- they contained phenomena or experiments that might be explained by a number of different ideas, all of which could be reasonably supported by some sort of evidence

Sub-sections in which groups spent a small percentage of their time engaged in SMD had one or more of the following characteristics:

- they tested and explored commonsense ideas, and so the students found the sub-sections neither helpful nor necessary
- they contained equipment or experimental procedures that distracted the groups from the primary sense-making purpose of the sub-section
- the primary sense-making purpose of the sub-section was no longer necessary because students had already settled on an idea that explained the phenomena in the sub-section

Result 5: Groups spent a large percentage of their time engaged in SMD in certain activities, and a very small percentage of their time engaged in SMD in others.

The same characteristics that supported or hindered SMD in activity sub-sections (see Result 4) also explained the existence or lack of SMD over entire activities. Additional factors that contributed to high sense-making percentages were:

- curriculum prompts that instructed groups to list the ideas for each group member and then come to agreement on which ideas might be the best (and why)
- exemplary peer tutoring and group interaction

<u>Result 6: The percentage of time that groups dedicated to sense-making discussion was lowest in cycle 3</u>.

The difference in SMD between curriculum cycles was not due to a difference in topics, but was instead due to the existence of one additional group member in group 2 (Jasper) during cycle 3. With Jasper in the group, group 2 contained two group members who were primarily interested in off-task behavior (Jasper, Roxanne), the result of which was that the off-task pair was nearly always successful in either (a) drawing one or both of the remaining members (Arthur, Sabrina) off-task, or (b) forcing Arthur and Sabrina to work silently, in isolation from the rest of the group.

<u>Result 7: The percentage of time dedicated to sense-making discussion was higher in group 1 than in group 2</u>.

Factors that contributed to the differences between groups 1 and 2 were that:

- Group 1 felt obligated to stay on-task, complete the activities as intended, and fill in all of the appropriate blanks. In general, group 2 did not feel these same obligations.
- Group 1 was more concerned than group 2 about understanding the underlying concepts of the CIPS curriculum.
- The leaders of group 1 (Darla and Lacey) encouraged the group to cooperate, complete the activities as intended, and evaluate and understand the concepts underlying the CIPS curriculum. The leader of group 2 (Roxanne) was almost always looking for an excuse to go off-task, and frequently succeeded in drawing other group members off-task.
- In group 2, one of the main contributors to sense-making discussion (Sabrina) maintained an expectation that verbal and written explanations should be simple and brief. The main contributors in group 1 did not share this same expectation.
- In group 2, there was friction between the two main contributors of sense-making statements. There was no friction between the main contributors in group 1.

<u>Result 8: In both groups, the amount and type of verbal sense-making per group member varied quite a bit</u>.

The factors that explained the sense-making distribution for each student were as follows:

Group 1: Darla. Darla was one of the leaders of group 1. Darla had a strong learning goal, and wouldn't settle for an idea that was unclear, problematic, or confusing; instead, she would involve the entire group in testing and modifying the troublesome idea. Also, Darla appeared to have a general interest in science. Darla's constant demand for elaboration, clarification, and explanation -- coupled with her propensity for peer tutoring -- is reflected in the large number of clarifications, explanations, and definitions/descriptions/connections in her distribution of sense-making instances.

Group 1: Lacey. Lacey was the other leader of group 1. Lacey was fairly conscientious about staying on-task and learning the underlying concepts of the CIPS curriculum. Lacey's verbal sensemaking was supported by Darla in that Darla was always available for SMD, and also because Darla would draw Lacey's attention back to the written materials when Lacey occasionally went off-task. Darla usually served as the ultimate authority and provider of guidance for group 1's SMD, which is why Lacey had fewer instances of defining, describing, and connecting than Darla.

Group 1: Grace. Grace's participation in group 1's sense-making discussions is not accurately reflected in her distribution of sense-making instances, which shows Grace's instances to be nearly equal in number and type to Darla's and Lacey's instances. It was either Darla or Lacey (or both) who typically drove the group's sense-making discussions, while Grace's sense-making statements were most often brief side comments on Lacey and Darla's back-and-forth conversation. Most of Grace's questions were not sense-making requests at all, but were instead requests for answers. Grace seemed not to be at a sufficient intellectual level and speed to contribute substantially to group 1's SMD.

Group 1: Porter. Porter's extremely small number of sense-making instances were due to the fact that (a) he preferred to work alone, and (b) when he did desire to engage in SMD with his group, his past history of working alone, coupled with his lack of communication and interpersonal skills, interfered with his attempts at SMD.

Group 2: Roxanne. Roxanne was the leader of group 2. Roxanne's sense-making statements were less frequent and less elaborate than the sense-making statements made by either Arthur or Sabrina. The reason that Roxanne had as many sense-making instances as she did is that there were a few isolated incidents where Roxanne engaged in extraordinary amounts of sense-making.

In general, Roxanne would constantly look for an excuse to go off-task, and often succeeded in drawing other group members off-task.

Group 2: Sabrina. Sabrina made more requests for sense-making than any other student. This was because she was the member of group 2 who felt most obligated to complete each activity's worksheets, keep the group on-task, and understand the activities' underlying concepts. Sabrina and Arthur were the two main contributors to group 2's SMD. Sabrina would typically initiate group 2's sense-making discussions by posing a question to Arthur, and then would sustain the discussions by considering Arthur's response and then resuming her questioning.

Group 2: Arthur. Like Darla, Arthur was a student with a fairly high interest in science. Generally, Arthur's role in group 2's SMD was not the initiator or director of sense-making, so much as a source of knowledge. Arthur's role is reflected in his distribution of sense-making instances, in that Arthur had more instances of verbal sense-making than any other member of his group. Many of these instances of verbal sense-making were direct responses to questions from Sabrina (more likely) or Roxanne (less likely); however, some instances were generating during Arthur's lectures to Roxanne and Sabrina on various scientific topics.

Group 2: Jasper. During the short time that he participated in the study, Jasper didn't engage in a single instance of verbal sense-making. Most of his time was spent off-task with Roxanne; when he was on-task, he tended to be loud, distracting, and verbally abusive. Jasper appeared to have little interest in understanding the principles of science or completing each day's activity. A common tactic was to let his groupmates do all of the work for an activity, and then copy the worksheet answers from a fellow group member or another member of the class.

Revisiting and Revising the Initial Factors List

At this point, I return to the factors framework for small-group SMD that was presented in Chapter 2. In that chapter, I summarized the factors that would likely affect small-group sense-making discussion. This list of factors consisted of factors that were known to affect nonverbal sense-making, as well as factors that were known to influence scientific discourse and group collaboration. These factors, seen in Table 2-3, were broadly sorted into four categories: personal factors, group factors, task factors, and contextual factors.

Table 2-3. Factors having a possible effect on the sense-making discussion in this study. (continued)

Personal factors		**Group factors**	
Student goals	Prior knowledge	Group expectations	Shared goals
Subject matter interest	Freedom from need	Leadership style	Collaborative skills
Cognitive incongruity	Metacognition	Student roles	
Comprehension: important and possible?			
Contextual factors		**Task factors**	
Classroom expectations	Role of teacher	Task goals	Science content
Assessments		Materials	Intrinsic motivation

Personal factors are those relatively stable intrinsic factors that one would normally associate with individual students (e.g., learning goals, interpersonal skills, and subject matter interest). Task factors reflect the various ways that the educational task influence discussion; such factors include the science content embodied in the task, the task goals, prompts in the curricular materials, and the degree to which the task is intrinsically motivating. Group factors, such as discussion roles and leadership styles, describe the ways that social interactions, group expectations, and group leadership affect the group's sense-making conversation. Contextual factors consist of the physical, organizational, and cultural aspects of the learning environment; examples of contextual factors include the physical layout of the classroom, class expectations, and the role of the teacher.

Ultimately, this broad framework of factors was utilized to validate whether these factors had a large effect on the group SMD in this particular study, and also to determine whether any additional factors influenced group SMD beyond the list of factors already documented. Therefore, the time has come to reorganize the Chapter 6 factors results (the factors that explain Result 1a, Result 1b, etc.) into group, task, personal, and contextual factors in order to determine:

- if the factors from the initial list did or did not support sense-making discussion in this study
- whether any new factors affecting sense-making discussion have been identified

Tables 6-11, 6-12, 6-13, and 6-14 list the group, personal, task, and contextual factors, respectively, that had a large effect on SMD in this study, and precisely what those effects were. New factors are italicized, and factors from the initial list that had little or no effect have been omitted, although hypotheses for why these factors did not significantly affect SMD are provided.

Table 6-11. Group factors affecting sense-making discussion in this study.

Factor	Ways in which this factor supported SMD	Ways in which this factor hindered SMD
Group expectations	Group 1 felt obligated to stay on-task, complete the activities as intended, and fill in the appropriate blanks in the curriculum materials. Since SMD was largely driven by the curriculum, these obligations promoted SMD.	Group 2 had members who did not feel obligated to stay on-task, complete the activities as intended, or fill in the appropriate blanks in the curriculum materials. In general, members of group 2 maintained an expectation that scientific explanations should be simple and brief.
Shared goals	Having two group leaders with learning goals kept group 1 focused on SMD.	Having two group members primarily concerned with off-task behavior suppressed group 2's SMD.
Leadership style	Group 1's leaders stressed cooperation, group learning, and the evaluation and reconstruction of troublesome ideas. Group 1 had a leader who experienced cognitive incongruity and was both willing and able to lead the group through extended sense-making discussions.	Group 2 had a leader whose primary goal was to be off-task, and often succeeded in drawing other group members off-task.
Collaborative skills	Group 1's leaders stressed cooperation, and were often successful in this regard.	Group 2 had personal friction between the main contributors to the group sense-making discussions. Some of group 2's members interrupted each other, did not respect each other's ideas, and were impolite. Group 1 had a member who preferred to work alone and lacked the communication and interpersonal skills needed to successfully communicate his ideas to the group.
Student roles	Group 1 had a member who actively participated in the activities, but copied answers from the other group members.	Group 2 had an active noncontributor who disrupted the conversation in the group.
Capacity for intra-group guidance	Group 1 had a knowledgeable member who could successfully engage in peer tutoring. Group 2 had a knowledgeable member who served a source of knowledge.	Group 1 sometimes reached a point where its members could not provide the intra-group guidance necessary to further develop the group ideas. The group became satisfied with the group ideas when they were still imprecise and inaccurate, even though there was experimental evidence available to clear up these imprecisions/inaccuracies.

Table 6-12. Personal factors affecting sense-making discussion in this study.

Factor	Ways in which this factor supported SMD	Ways in which this factor hindered SMD
Student goals	Strong learning goals supported Darla and Lacey's SMD. Darla wouldn't settle for ideas that were unclear, problematic, or confusing.	Social and off-task goals sometimes kept members of group 2 off-task. Grace (group 1) and Roxanne (group 2) were primarily concerned with getting answers; they regularly copied answers from other group members. Roxanne, especially, had little interest in understanding the principles of science. Students occasionally saw guesswork as acceptable.
Prior knowledge	Arthur's extracurricular knowledge of science allowed him to be a source of knowledge for other group members.	
Cognitive incongruity		Some activities couldn't provoke cognitive incongruity in a single group member.
Metacognition	Darla was capable of (and intensely interested in) monitoring problems with her ideas.	
Comprehension: important and possible?	Darla and Lacey had learning goals, and so saw comprehension as important.	Many members of group 2 did not see comprehension as important.
Subject matter interest	Darla and Arthur had a general interest in science. Darla drew on this interest to conduct SMD in her group, but Arthur only used this interest in his capacity as a source of knowledge for other group members.	
Intellectual capacity		Grace was at an intellectual level far below the other members of her group; this kept her from contributing substantially to group 1's SMD.

One of Hatano's requirements for sense-making is that students must view comprehension as both important and possible, but it is interesting that the students in groups 1 and 2 only had difficulty with the former: thinking of comprehension as important. At various times, students were off-task over the course of the study, and for a variety of different reasons: social goals, distraction from a fellow group member, and so forth. These off-task episodes illustrated that students sometimes preferred to put their efforts into socializing rather than into comprehension. But it never appeared that the students refrained from SMD because they lacked confidence in their abilities to construct scientific explanations.

One possible reason for the unseen "confidence" effect is not so much that the students had extremely positive views of their ability to do science, but instead that they were more neutral on the matter, meaning that the middle school students in this study were neither positive nor negative about their scientific abilities. Another possible explanation is that, in order for students to lack confidence in their ability to explain, they must first realize that scientific explanations can be very complicated, and therefore can be quite challenging to construct. It is only when students realize that scientific explanations should be complete, precise, sufficiently detailed, and potentially generalizable that they are able to recognize that they may or may not be up to the task of constructing a scientific explanation. A good fraction of students in this study may not have viewed explanations as necessarily precise or detailed, and so the opportunity for these students to realize that they may not be able to construct these "complete" scientific explanations likely never arose.

Table 6-13. Task factors affecting sense-making discussion in this study.

Factor	Ways in which this factor supported SMD	Ways in which this factor hindered SMD
Task goals	Many sub-sections were explicitly directed at sense-making, and -- as expected -- students spent more time on SMD in these sections.	
Science content		Predictions or experiments that tested unsurprising, commonsense ideas led to very little SMD. Students settled on explanations early in their discussion of a phenomenon if they did not have enough knowledge of (or experience with) the phenomenon to support their ideas with reasonable evidence. Students became distracted from the primary sense-making focus of certain activities when they engaged in SMD about conceptual clarifications that were not addressed by the curriculum.
Materials	Energy and force diagrams provided the conceptual scaffolding necessary for students to clarify, differentiate, and discuss related concepts and processes. Force diagrams required students to be precise in their thinking, which pushed them to a level of SMD that might otherwise not have been possible. Prompts for students to record each other's ideas, and then agree on a "best" idea and reasoning, were successful in promoting discussion.	Some activities contained equipment or experimental procedures that distracted the groups from the primary sense-making purpose of the activity.
Intrinsic motivation	Activities that allowed students to take creative control and were both novel and personally relevant gave rise to prolonged SMD because the students actually enjoyed the SMD in these activities.	

Table 6-14. Contextual factors affecting sense-making discussion in this study.

Factor	Ways in which this factor supported SMD	Ways in which this factor hindered SMD
Role of teacher		Occasionally, the teacher did not emphasize that students should be sure to complete the activity sections primarily dedicated to sense-making, and so students would skip or spend little time in these sections. At times, the teacher hinted at (and sometimes even told) the class the correct benchmark ideas early in the idea development cycle.
Class expectations	(effects unclear)	(effects unclear)
Time		Sometimes there was not enough class time for students to complete the activity sections primarily dedicated to sense-making.
External guidance		There was no mechanism for groups to get external guidance in those cases where students became satisfied almost immediately with their imprecise or incorrect group ideas.
Awareness of curriculum structure		Teachers and students were unclear as to which curriculum sections were to be done in groups and which sections were to be done individually, and so some group sections were done individually (and vice versa).
Awareness and valuing of curriculum goals		Some students did not value the ongoing revisiting and redevelopment of their ideas, which was the primary goal of the CIPS curriculum

The more interesting aspects of Table 6-14 are that the effects of class expectations were unclear and that the "assessments" factor is missing.

It is curious that assessments had little effect, especially in light of the fact that the CIPS Power Drive assessments (or "tests", in the words of the teacher) were fully available to the students in the curriculum materials.

Students were well aware that the last activity of each cycle was their "test" (the teacher told them so), but -- despite this fact -- not once did I see a student in either group look ahead to see what the cycle Power Drive might be. I suggest that students' lack of concern for their assessments was derived from four sources: (1) some students weren't motivated by grades, most likely because they didn't much care what grade they received, (2) it wasn't made explicit to the students how the Power Drive assessments fit into their overall grade, (3) because student work consisted entirely of in-class experimental inquiry aimed at the development of ideas (rather than the understanding of "correct" ideas), it was difficult for students to explicitly prepare for the Power Drive assessments, and (4) if the class as a whole fared poorly enough on an assessment, the class was allowed to go through the assessment a second time. Combined, these four factors meant that students were not terribly concerned about learning the material with the intent of doing well on the CIPS Power Drive assessments.

The effects of class expectations on small-group SMD were unclear because, although I recorded and analyzed any direct effects that whole-class activities or discussions had on the small-group SMD in this study, the effects of class expectations were more indirect, and were therefore much more difficult to identify. As an example, two characteristic of the whole-class discussions/presentations in this study were that: (1) very few students participated in whole-class discussions, and (2) presenting students often could not defend or explain the group ideas, and frequently would refer teacher or student questions to another group member.[17] The likely effect of both (1) and (2) is that, over time, students knew (expected) that they didn't have to be prepared for whole-class presentations or discussions; consequently, students didn't feel the need to engage in group sense-making, or even pay attention to their group's activities, to the extent that they should have. I use the phrase "likely effect" here because both the expectation itself ("I won't have to participate significantly in class discussions") and the effect of that supposed expectation (less effort/attention in group activities) are difficult to prove. There is little chance that students will voice this type of expectation, and the lack of effort in group activities is more easily attributable to personal, group, or task factors than to class expectations. Therefore, because of the difficulty in identifying class expectations and their effects, I cannot state with confidence any of these expectations or effects in Chapter 6.

[17] These characteristics are explored further in Chapter 7.