Chapter 6: The Focus of Reflection

Introduction and Rationale

Chapter 5 indicated that prompts for reflection do, in some cases, make a difference in the ways that students succeed. What accounts for these differences? The pilot work indicated that students respond to similar prompts in varied ways. One hypothesis might be that the focus of students' reflection plays an important role in determining how they succeed given a certain type of reflection prompt. How do students reflect? What kinds of reflection are most effective in fostering learning? The current chapter attempts to address these questions and others that are related.

In this chapter I describe how students interpret and make use of the prompts they were given, allowing a comparison at the level of the two conditions. Further analysis allows us to identify predictors of the most effective learning combinations—that is, what is the most effective way to combine instructional strategies with learner propensities toward certain types of reflection to foster knowledge integration most consistently.

Methods

In addition to considering the prompt type (i.e., directed or generic), the analysis discussed in this chapter investigates the effects of prompt placement, too. By prompt placement, I mean where the reflection prompt came with respect to the activity itself. As discussed in Chapter 4, the prompts in each condition occurred at exactly the same points in the project. The generic prompts were indistinguishable from one another, regardless of placement. However, the 11 directed prompts belonged to three distinct categories based largely on their placement with regard to the activity. Specifically, 6 were termed "Thinking Ahead," 4 were termed "Checking Our Understanding," and 1 was termed "Thinking Back."

Thinking Ahead prompts encouraged planning for future activities and typically were located before a particular activity. Checking Our Understanding prompts asked students to monitor their understanding, and were typically located after an activity. The lone Thinking Back prompt occurred at the very end of the project, and encouraged students to reflect back on how they worked on the project and assess their strengths and weaknesses. For the directed prompts, then, we can investigate the differences in the reflection elicited for the three orientations of prompts.

Data Sources and Outcome Measures

The primary data sources for investigating the role of the focus of reflection in students' learning include the students' prompt responses and their projects themselves. (The sources of data within the projects have been discussed in some depth in Chapter 5.) Other sources of data include quizzes and tests as well as in-depth interviews with a select group of students. As in Chapter 5, qualitative data are interspersed with the quantitative results. Outcome measures are associated with these data sources, and will be addressed in turn.
Reflection Prompt Responses

Responses to reflection prompts were assessed on their length and the focus of their reflection.

The degree of elaboration was assessed most simply by counting the number of words in the response. That is, the total number of words in the note was tallied, and then the number of words in the prompt itself was subtracted off, leaving only the words in the response itself to be counted. Another, more complex measure of the degree of elaboration was the number of comments students made. This can also be thought of as the number of ideas included in the response. For example, in response to a generic prompt, one pair of students wrote a response including 4 comments:

What we are thinking about now is... the article that we received in class. It is talking about staying cool in the summer by wearing a white shirt. In this article we notice that claim three is incorrect. Nothing can stay the same temperature all the time, and especially not metal.

The 4 comments include "the article that we received in class," "It is talking about staying cool in the summer by wearing a white shirt," "claim three is incorrect," and "Nothing can stay the same temperature all the time, and especially not metal." This example, then, would receive scores of 4 comments and 44 words (51 words total minus 7 words in the prompt). Note that in this example, all the comments were written by the students as separate sentences; this is not necessarily the case.

The number of words and number of comments measures combine to create the third measure of elaboration, which was the number of words per comment. This last measure allowed assessment of elaboration of each idea, in addition to the total elaboration (number of words) and total ideas (number of comments). It is developed using the two individual counts, rather than counting the words in each comment, since often the decision of where a comment actually starts and stops may be ambiguous. The example above would receive a score of 11 words per comment; on average, each comment in the example included 11 words.

Students' reflection focus was also assessed in this aspect of the study. Reflection focuses were broken into 5 major codes, as discussed in Table 6–1 below. The data contained responses to each type of prompt showing each of the major codes.

<table>
<thead>
<tr>
<th>Major Codes</th>
<th>Sub-Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflection on Actions: involves thinking about what ways of behaving are most appropriate, as well as contemplating very general goals</td>
<td>Schoolish Social General goals</td>
</tr>
<tr>
<td>Reflection on Activities: involves thinking specifically about the project</td>
<td>Project-specific goals</td>
</tr>
<tr>
<td>Reflection on Project Ideas: involves thinking about the ideas presented in the project itself</td>
<td>Content Critique</td>
</tr>
<tr>
<td>Reflection on Knowledge: indicates a focus on monitoring or improving one's understanding</td>
<td>Thought Knowledge Integration</td>
</tr>
<tr>
<td>Non-Reflective: includes responses of students who say they have &quot;no problem&quot; and those of students who say their thoughts haven't changed since the last prompt</td>
<td>&quot;No Problem&quot; &quot;Same as before&quot;</td>
</tr>
<tr>
<td>Other: includes those responses that are uncodable in the previous scheme</td>
<td>Other</td>
</tr>
</tbody>
</table>

Table 6–1: Coding Reflection in Prompt Responses
The major codes were further broken into sub-codes. Examples of these sub-codes are given in Table 6–2. For ease of interpretation, Table 6–2 includes only a single comment of each type in the exemplars given. However, generally, students responded with multiple comments, and often those comments received different sub-codes. (For instance, in the example given previously to show the way responses were broken into comments, the comments would be coded, respectively, as focused on activities, content, critique, and knowledge.) Also, Table 6–2 gives examples mainly from directed prompts, but the same coding scheme was used for the generic prompts, as well. Only in the generic prompt condition did students write responses requiring the "Same" sub-code, so its exemplar is a generic prompt response.

<table>
<thead>
<tr>
<th>Sub-Codes</th>
<th>Example of Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFLECTION ON ACTIONS</td>
<td></td>
</tr>
<tr>
<td>Schoolish</td>
<td>To do a good job on this project, we need to… double check our answers before we turn them in</td>
</tr>
<tr>
<td>Social</td>
<td>To do a good job on this project, we need to… talk and think through this project together</td>
</tr>
<tr>
<td>General goals</td>
<td>To do a good job on this project, we need to… look at all the evidence</td>
</tr>
<tr>
<td>REFLECTION ON ACTIVITIES</td>
<td></td>
</tr>
<tr>
<td>Project goals</td>
<td>To do a good job on this project, we need to… make a critique of at least three of the evidence</td>
</tr>
<tr>
<td>REFLECTION ON PROJECT IDEAS</td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>Pieces of evidence we didn't understand very well included… why the metal and marble benches are colder than the wooden benches</td>
</tr>
<tr>
<td>Critique</td>
<td>Pieces of evidence we didn't understand very well included… there was not a distinct scientific basis to the [benches] experiment, therefore it was not absolutely believable</td>
</tr>
<tr>
<td>REFLECTION ON KNOWLEDGE</td>
<td></td>
</tr>
<tr>
<td>Thought</td>
<td>When we critique evidence, we need to… understand the evidence</td>
</tr>
<tr>
<td>Knowledge Integration</td>
<td>When we critique evidence, we need to… think back to experiments we did before</td>
</tr>
<tr>
<td>NON-REFLECTIVE</td>
<td></td>
</tr>
<tr>
<td>&quot;No Problem&quot;</td>
<td>Pieces of evidence we didn't understand very well included… all of them were very clear to both of us</td>
</tr>
<tr>
<td>&quot;Same as before&quot;</td>
<td>What we’re thinking about now is… basically the same as the beginning.</td>
</tr>
<tr>
<td>OTHER</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>When we critique evidence, we need to… keep in mind that what we say to this company is what is going to keep its readers</td>
</tr>
</tbody>
</table>

Table 6–2: Exemplars of Reflection Types

These codes break down loosely into the metacognition/sense-making taxonomy for reflection: students who focus on actions, activities, and knowledge are being metacognitive (to varying degrees) and those who focus on project ideas are engaging in sense-making.

Other Outcome Measures

As in the analysis discussed in Chapter 5, other primary outcome measures for this aspect of the study are associated with the students’ projects—in particular, their claim notes and letters to the editor. As in Chapter 5, these outcome measures include primarily measures of coherence of ideas, overall critique quality, and guidelines quality, as well as the overall
E. A. Davis

project score. These measures were all discussed in depth in Chapter 5 and will not be revisited in depth here. In the current aspect of the study, these project quality measures are linked to the reflection prompt measures discussed above (elaboration and focus of reflection).

Analyses

The analyses used to investigate the role reflection plays in student performance on the project include t-tests and factorial and repeated measure ANOVAs to compare the focuses of reflection elicited by the two conditions. Correlations between reflection focus and success on various quality measures elucidate productive and less productive focuses for reflection. The correlational analyses will be used as a foundation for multiple regressions combining several factors; those synthesizing analyses are discussed in Chapter 8.

Most of the analyses compare directed and generic prompts. Some, however, further break the analysis down by prompt placement (i.e., Thinking Ahead, Checking Our Understanding, and Thinking Back for directed prompts). Where appropriate, the Thinking Back prompt has been eliminated from the analysis because of its singular nature in the intervention. These analyses will be noted as such.

Reflection Prompts and Project Quality Results

The results are separated into two major areas. The first concentrates on the actual responses to the reflection prompts, and the second links students' focuses of reflection to their success on the project. To preview my findings, I show that a non-reflective type of response is relatively prominent for the directed prompts as compared to the generic prompts, and that students who do not take advantage of the opportunities for reflection provided by prompts are generally less successful on the project than are other, more reflective students. I hypothesize that these non-reflectors lack opportunities other students have to identify weaknesses in their current knowledge.

Reflection in Response to Reflection Prompts

Recall that students' responses to the reflection prompts are coded on the basis of elaboration of reflection and focus of reflection.

Elaboration

How did the prompt conditions differ in the amount of reflection elicited? That is, how did students' responses to directed and generic prompts differ in the degree of elaboration? (Elaboration is used as a measure of the amount of reflection—though we will see that not all reflection is equally productive.)

The two conditions did not differ in the number of words students wrote in response to the two types of prompts. The overall mean number of words was 23.2 (SD = 9.8; see Table 6–3). However, there were significant differences in the number of comments and in the number of words per comment elicited by the two types of prompts. On average, students
wrote significantly more comments in response to directed prompts than to generic prompts ($F[1, 89] = 6.264, p = .0142$) but elaborated their comments significantly more when responding to generic prompts ($F[1, 89] = 26.741, p < .0001$; see Table 6–3).

<table>
<thead>
<tr>
<th></th>
<th>Overall mean (SD)</th>
<th>Directed Prompt condition mean (SD)</th>
<th>Generic Prompt condition mean (SD)</th>
<th>Significant difference between conditions?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Words</td>
<td>23.175 (9.786)</td>
<td>22.648 (9.633)</td>
<td>23.669 (10.006)</td>
<td>no</td>
</tr>
<tr>
<td>Number Comments</td>
<td>2.164 (0.674)</td>
<td>2.341 (0.638)</td>
<td>1.998 (0.670)</td>
<td>yes</td>
</tr>
<tr>
<td>Number Words/Comment</td>
<td>10.557 (2.284)</td>
<td>9.428 (2.098)</td>
<td>11.613 (1.933)</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 6–3: Degree of Elaboration of Reflection Prompt Responses (N = 91)

These results indicate that in their immediate responses, directed prompts elicit a greater number of ideas, but that generic prompts encourage greater elaboration of those ideas.

Students could be terse in response to either kind of prompt. For example, here are a few examples of students’ responses to directed and generic prompts:

- Pieces of evidence we didn't understand very well included…
  - S124 & S127: ...a lot of the science theories.
  - S306: ...none.
  - S302 & S330: We understood all of the pieces of evidence.
  - Our ideas right now are…
  - S515 & S525: … that there should be more information and evidence in this project.
  - S516 & S532: … that the evidence they gave was pretty true.

Compare those examples with these examples of elaborated responses:

- Pieces of evidence we didn't understand very well included...
  - S106 & S110: ... the chemical or 'anti-heat' shirt because they didn't really explain about the shirt was made out of or evidence that the shirt actually worked. We also didn't understand the small and large rooms one because it didn't give their original temperature's of the room and their were three people in the small room and only one in the large room so that there was more body heat in the small room to begin with so that they should have the same amount of people in each room.
  - Our ideas right now are …
  - S505 & S507: … that this newspaper should conduct better experiments and check their claims with real scientists before they print an article. They should also not make assumptions, but come to real conclusions from well conducted experiments.
E. A. Davis

S712 & S715: … that the article could be changed to sound right, and actually be correct, if the authors looked over the evidence closer before making their claims, and looking at the creditability of the evidence.

Clearly, students who elaborate their responses often are reflecting more deeply about various kinds of ideas—or are articulating their reflection more completely.

**Reflection Focus**

How did students differ in the reflection they engaged in as a result of the two types of prompts? How does prompt placement affect their responses?

Factorial ANOVAs run on each possible focus of reflection indicate significant differences in students' reflection in response to directed and generic prompts. (Refer back to Table 6–2 for examples of each focus of reflection.) Students who received directed prompts focused on actions significantly more often ($F[1, 89] = 14.421, p = .0003$) and were non-reflective significantly more often ($F[1, 89] = 14.000, p = .0003$). Those who received generic prompts, on the other hand, were significantly more likely to focus on project ideas ($F[1, 89] = 10.517, p = .0017$) and on knowledge ($F[1, 89] = 4.111, p = .0456$). There was no difference in the amount of reflection focused on activities ($F[1, 89] = 0.214, p = .6447$). Figure 6–1 presents the relative frequencies of each type of reflection.

![Figure 6–1: Reflection Focus for Directed and Generic Prompts](chart)

These data indicate, first, that both conditions elicit mostly reflection focused on project-specific goals (activities), and that there is no difference in the degree to which student focus in this direction in the two conditions. Secondly, these data imply that generic prompts may elicit more reflection on higher-level focuses, because students in the generic prompt condition focused more on content, critiquing, and knowledge, whereas students in the directed prompt condition focused instead on schoolish, social, or general goals or, alternatively, were non-reflective. (Refer back to Tables 6–1 and 6–2 for the codes and sub-codes of reflection focus.)
The following examples are of typical responses to directed prompts. The first focuses on actions and the second is non-reflective.

S102 & S108: To do a good job on this project, we need to read our evidence thoroughly and then take good notes.

S317 & S326: Claims in the article we didn't understand very well included none, we understood them all.

A metacognitive response to a generic prompt is quoted here:

S205 & S218: Our thoughts now are about things that we have already learned this year in science as well in previous years.

While these examples clearly comprise too small a sample to show the full range of student responses, they do give a sense of the ways in which students reflected in response to the two levels of specificity of reflection prompts.

What effect did prompt placement within the project have? We would hypothesize that "placement" (i.e., Thinking Ahead, Checking Our Understanding, and Thinking Back) would have a strong effect for directed prompts, since prompts differed in terms of orientation. We would expect no effect for generic prompts, since these were virtually identical across the project.

Indeed, for most focuses, paired t-tests indicate significant differences in reflection focus for Thinking Ahead and Checking Our Understanding directed prompts, but no significant differences for corresponding generic prompts. Students in the directed prompt condition reflected on actions (t[43] = 9.333, p < .0001) and activities (t[43] = 16.352, p < .0001) significantly more often when responding to Thinking Ahead prompts than when responding to Checking Our Understanding prompts, but focused on project ideas (t[43] = 12.617, p < .0001) and were non-reflective (t[43] = 5.365, p < .0001) significantly more often for Checking Our Understanding prompts than for Thinking Ahead prompts. There were no significant differences in the amount students focused on knowledge in response to prompts with different orientations. The only focus for which placement made a difference for generic prompts was that of being non-reflective; as with directed prompts, students responded to generic prompts at the Checking Our Understanding locations (typically immediately after the activity) in a non-reflective way more often than they did so for prompts at the Thinking Ahead locations (t[46] = 2.871, p = .0062), perhaps because students felt they had "just said" what they were thinking. Figure 6–2 shows these differences, and Table 6–4 provides the proportion of reflection with each focus.
These results indicate that, as expected, Thinking Ahead and Checking Our Understanding prompts elicited quite different kinds of reflection. Students tended to respond to Thinking Ahead prompts with reflection focused on activities (almost 60% of the time), knowledge (20%), and actions (18%), whereas when they received prompts asking them to monitor their understanding, they focused mainly on project ideas (about 55% of the time), knowledge (16%), and were non-reflective (21%). That is, directed prompt students' plans tended to focus on project-specific goals, knowledge, and schoolish sorts of thinking. Their monitoring tended to focus instead on content, critiquing, and knowledge—as well as showing a lack of reflection. For generic prompts, as expected, students typically responded similarly regardless of the location of the prompt. Generic prompt responses focused largely on activities, project ideas, and knowledge.

**Prompt Responses and Project Characteristics**

We have seen the range of students' reflection and how it differs given different prompting experiences. How does that reflection relate to students' work on the project?
Initial analyses related qualities of students' reflection to their overall project scores, as determined using a holistic scoring system based on the degree of completion and conceptual quality and as discussed in Chapter 5. Later analyses are more specific, relating reflection to project quality measures such as number and type of cites made, the overall critique quality, the coherence of students' specific ideas, and the guidelines quality. We look first at how elaboration of reflection prompt responses is related to these characteristics.

**Elaboration and Project Success**

I measured elaboration by the number of words in the prompt response—a measure for which there existed no differences for the two conditions—and the number of comments—which was higher for the directed prompt condition. When we look at more specific relationships between elaboration of reflection prompt responses and project quality, we find that elaboration of generic prompts was not significantly correlated with any of the three main project quality measures (namely, overall critique quality, coherence of scientific ideas, and guidelines quality). For students in the directed prompt condition, on the other hand, elaboration of directed prompt responses was significantly positively related to overall critique quality—especially for Checking Our Understanding prompts ($r = .435, p = .0032$ for number of words; $r = .308, p = .0444$ for number of comments). It appears that elaborating on the quality of one's understanding is a good predictor of successful critiquing of evidence and claims, perhaps because students have identified, in the context of the Checking Our Understanding prompts, confusing or misleading aspects of the evidence and claims, pointing to aspects worthy of critique. They can then delineate ways in which the evidence or claims could be improved. (The generic prompts lack that evaluative context and elaboration of them may not be as useful.)

Elaboration of directed prompts was not significantly related to either of the other two project quality measures.

Here is an example of an elaborated Checking Our Understanding response for a pair that ended up doing well (in the top quartile) on overall critique quality:

S107 & S121: Claims in the article we didn't understand very well included the fact that if you were in a room with lots of marble and metal in it, you would stay cooler. Perhaps if you were touching something marble or metal, you would stay cooler, but I'm not so sure what the results would be in a room. Also, the anti-heat shirt confuses me. What chemical is this shirt dipped in and how can it affect one's body temperature? How does it work? Does such a thing really exist?

These students wrote a good critique of claim 3 (the thermal equilibrium claim), which they started concentrating on in this first Checking Our Understanding note. The following is an excerpt from the same pair's letter:

For the claim "some materials are naturally cold" we thought this claim was semi-valid. It didn't really explain what makes marble and metal benches colder. It also didn't really prove that if you stayed in a room with lots of marble and metal in it, you would stay cooler. When improving this claim, we think that you should include evidence proving how the marble and metal benches would be colder and how it would make the room cooler.

The scientific ideas for this evidence involve the fact that even though the room temperature feels the same, the material of the object affects and determines how it functions (absorption, attraction, etc.) in a certain climate. The methods of this evidence show that even if some surfaces feel different than others, their temperature remains the same. It sometimes depends on how the object will react to
different amounts of temperature. The credibility for this evidence is very high because you can relate to it very easily, as with the use of park benches of different materials in the same climate. This is also very believable because the same thing happens to me at the beach (the sand is hot and the grass and towel are cool) as well as with benches of different materials. We think that the evidence is very useful because it shows how materials have different feeling temperatures even in the same climate and temperature. Picnic benches made of wood feel much more different or warmer than a steel bench, even if they are in the same climate and temperature.

These students critique the evidence on the basis of science, methods, and credibility, pulling in real-life experiences and citing a principle of thermal equilibrium. They identify materials as having different thermal characteristics; they do not link the idea of rate of heat flow, though, relying instead on materials "reacting differently to temperature." Nonetheless, their critique is quite impressive and may be linked in part to their elaboration of Checking Our Understanding notes—their average number of words used in responses to Checking Our Understanding prompts was 50.5, as compared to the overall average of 23.2.

Reflection Focus and Project Success

How is students' focus of reflection related to their success on their projects? Students' focus on aspects of the project ideas and the degree to which they write uncodable or non-reflective responses were all related to overall success and performance on the more specific project quality measures.

Interestingly, focusing on project ideas and on content (which makes up one aspect of the project ideas major code) had opposite effects for the two prompt conditions. For the directed prompts, focusing on project ideas was significantly positively related to overall project score—in particular, focusing on project ideas in response to Checking Our Understanding prompts was positively related (r[44] = .311, p = .0391 for all directed prompts; r[44] = .304, p = .0444 for Checking Our Understanding prompts in particular). However, for the generic prompts, focusing on content was negatively related to overall project score (r[47] = −.302, p = .0389).

What, specifically, was a project ideas- or content-focus related to? For the directed prompts, focusing on content or project ideas in response, in particular, to Checking Our Understanding prompts was significantly positively related to developing coherent ideas (r[43] = .317, p = .0377 for content focus; r[43] = .335, p = .0274 for project ideas focus). These focuses were not significantly related to the other project quality measures for directed prompts. Focusing on content in response to a prompt intended to elicit monitoring may help students develop coherent ideas.

This may be related to the opportunity that directed prompts afford for thinking more deeply about the content. It is easier to say, "Right now I'm thinking about energy conversion" without taking any action to further one's understanding of energy conversion than it is to say "I don't understand energy conversion" without taking action. Admitting a lack of understanding affords the opportunity, for many people, to work to better understand the topic by identifying places where they could link, distinguish, or reorganize their ideas. Some examples of typical content-focused Checking Our Understanding prompts show the kinds of extended reflection that directed prompts can afford:
E. A. Davis

S418 & S408: Pieces of evidence we didn't understand very well included the Goldilocks and the Three Benches. There was not a distinct scientific basis to the experiment therefore it was not absolutely believable. The evidence was basically based on substantial evidence.

S428 & S403: Claims in the article we didn't understand very well included the claim about the marble and metal benches. It seems to me that when you sit in a metal bench that has been sitting the sun it would be very hot. Also the claim about the shirt doesn't really give me enough evidence about the shirt or where to get for me to believe if it actually work.

Both of these pairs ended up with high (top quartile) scores in coherence. The first pair goes on to critique the content (the "Benches" evidence; see Appendix F), and the second pair links their confusion about the claim associated with the "Benches" evidence to an experience of their own and distinguishes between the temperature of metal in a hot room versus in a cold one. They assess their understanding (i.e., they identify what they are confused by) and then extend their reflection in various, potentially productive ways. Contrast those examples with this typical content-focused response to a generic prompt:

S221 & S217: Right now, we're thinking about the energy conversion principles that indicate that black attracts heat.

This pair developed much less integrated, coherent ideas than did the other pairs discussed above. One might hypothesize that they did not ever assess that they did not fully understand the energy conversion claim; they simply stated that they were thinking about it. Thus, they did not identify a weakness in their knowledge, and missed an opportunity to distinguish (in this case) between attracting and absorbing, and between heat and light. We see that the positive (or negative) effects of reflection (or lack thereof) in response to prompts propagate through the students' work on the projects.

In fact, for the generic prompts, there exist significant negative relationships between a focus on content and the measures of overall critique quality (r[46] = –.380, p = .0087), coherence of ideas (r[47] = –.388, p = .0066), and guidelines quality (r[47] = –.298, p = .0418). Additionally, guidelines quality also bore a negative relationship to focusing on project ideas (that is, both content and critiquing; r[47] = –.329, p = .0233). These negative relationships between a content focus in response to generic prompts and success on all aspects of the project may indicate that focusing on content requires an assessment or evaluative component in order to be productive. Students who focused on content without the context of monitoring their understanding were likely to develop incoherent ideas, to fail in critiquing evidence and claims, and to not be able to abstract out from the act of critiquing to develop guidelines for critiquing. On the other hand, those who focused on content in response to Checking Our Understanding prompts were likely to develop more coherent ideas.

I also identified a strong negative relationship between writing uncodable responses to directed prompts and project success, though only about 2% of the responses to either type of prompt were uncodable. There was no relationship between writing uncodable responses to generic prompts and any of the specific quality measures. However, for the directed prompts, writing uncodable responses was significantly related to both overall critique quality (r[43] = -.351, p = .0204 for uncodable responses overall) and guidelines quality (r[43] = –.517, p = .0003 for uncodable responses overall; r[43] = –.418, p = .0048 for uncodable Thinking Ahead responses). While writing uncodable responses to directed prompts was not correlated with (in)coherence of ideas, it was related to poor critiquing and poor guidelines for critiquing. Directed prompts may allow some students to
create highly useful plans for critiquing, and to thus develop a better understanding of critiquing—but for students whose own reflection does not map on to the “direction” of directed prompts, the prompts may instead derail them. That development of an understanding of what it means to critique may then lag behind that of other students. As with students who focus on content in response to generic prompts, these students miss an opportunity to identify weaknesses in their understanding of which other students take advantage. Generic prompts, by allowing students to determine their own reflection direction, may allow all students to benefit from reflection (even if that reflection is not strictly interpretable).

A similar relationship between how students responded to reflection prompts and their overall success on the project was found in the degree to which students are non-reflective. There existed no significant correlations between being non-reflective and project success for the generic prompts. However, in the directed prompts, we see a different story. Students who were unreflective in response to Checking Our Understanding prompts were significantly less likely to succeed on the project (based on their overall project scores) than were those who were more reflective ($r_{44} = -.481, p = .0008$).

This relationship for the non-reflective students who received directed prompts manifested itself mainly in the quality of students’ critiques. There existed a negative relationship between being unreflective in response to Checking Our Understanding prompts and overall critique quality ($r_{43} = -.302, p = .0490$). Students who perceive themselves as understanding all the ideas may be less likely to do a good job at critiquing the evidence and claims, perhaps because they are less likely to identify problems in those evidence and claims themselves, as well. (Students who were non-reflective in response to Checking Our Understanding prompts tended also not to elaborate their responses to those prompts. For example, a common non-reflective response to the prompt “Pieces of evidence we didn't understand very well included...” was “none.” Both being non-reflective and writing lengthy responses to reflection prompts were positively correlated with critiquing; we cannot yet determine the causality of these relationships.)

Here are two examples of non-reflective (and relatively terse) responses to Checking Our Understanding prompts for a pair that did poorly at critiquing:

S102 & S108: Pieces of evidence we didn't understand very well included none of them. They all make sense. Just come of them are not really useful.

In thinking about how these ideas all fit together, we're confused about nothing at all because we understand it all.

In these students' evidence and claim notes, when critiquing the "Bikes at Night" evidence, they wrote summaries rather than critiques. For example, in "critiquing" the methods, they wrote:

The methods for this evidence were that they got two people. One with a white shirt and one with a black shirt. They tried to show that the person with the white shirt could be seen better than the guy with the black.

When critiquing the third claim, these students wrote:

In thinking about the validity of this claim, we think that this claim is very valid because it has good proof to support it and it is accurate and makes sense.
Clearly these students did not assess their own understanding as problematic. Perhaps they lacked an understanding of the purpose of critiquing (as exemplified by the first critique example), or perhaps they were equally unlikely to assess a claim or piece of evidence as flawed as they were to see a flaw in their own understanding of the claim or evidence. As with the students who wrote cryptic responses, they missed an opportunity to identify weaknesses in their own knowledge, or places where they should link or distinguish ideas.

**Summary and Implications**

Both kinds of prompts appear to have strengths and weaknesses, but as we saw in Chapter 5, generic prompts look a bit more promising. However, as we also saw in Chapter 5, they are not without their problems.

Both kinds of prompts elicit elaboration on the part of the students. Generic prompts, however, appear to elicit more conceptual and abstract reflection, in that students who receive generic prompts focus more on project ideas and knowledge, whereas those who receive directed prompts focus on more logistical ideas like actions, or are truly non-reflective. This reflection focus, however, only plays out as success (or lack thereof) on the project in limited ways. For example, students who focused on content in response to directed prompts were likely to develop a coherent understanding of the science. However, students who focused on content in response to generic prompts—a focus particularly prevalent for generic prompts—tended to do poorly on all measures of project quality. Generic prompts may discourage some students from identifying weaknesses in their current knowledge by allowing them to report on content rather than reflecting on their understanding of it.

Directed prompts may have more severe negative effects. Students who wrote uncodable or non-reflective responses to directed prompts tended to do worse on the project than those who took more productive routes for their reflection. Recall, too, that directed prompts—especially Checking Our Understanding prompts—elicited significantly more non-reflective responses than did generic prompts. This lack of reflection has serious consequences for students' success. Students who misdiagnose their understanding or even reflect in cryptic ways may eschew opportunities to distinguish among their ideas, link those ideas, and identify weaknesses in their knowledge.

It may be that directed prompts enable some students to succeed (if, for example, they focus on project ideas in response to Checking Our Understanding prompts) at the same time as they enable other students to fail (if, for example, they improperly monitor their understanding by responding to Checking Our Understanding prompts in a non-reflective way). Essentially, it seems that directed prompts may help some students but hurt others, perhaps based on their self-monitoring skills. Success with directed prompts appears based in large part on the content of their reflection; their reflection may elicit or stunt the processes of knowledge integration. Those who do not actually reflect are much less successful.

With the exception of the negative effects of focusing on content, the actual orientation of students' reflection in response to generic prompts seems less important when related to project quality. Generic prompts allow—in fact, force—students to reflect in their own 'default' ways. These defaults may be, for many students, more useful than our best-intentioned direction; most students appear capable of productively taking charge of their
own reflection. In particular, generic prompts do not generally enable students to be “cognitive economists”—students who do the minimal amount of cognitive work (Linn, Songer, & Eylon, 1996)—as do directed prompts. If they manage to identify weaknesses in their current knowledge, students responding to generic prompts may begin to distinguish among ideas and link those ideas both within their responses to the generic prompts themselves, and, more often, in aspects of the project. Since generic prompts appear to help students expand their repertoire (through eliciting principles and other cites) and identify weaknesses in their existing knowledge (by forcing some kind of reflection), students in the generic prompt condition appear to have more opportunities to integrate their knowledge. And in fact, students responding to generic prompts developed a more coherent, integrated understanding of the science concepts than did other students.

How do these results differ when we view them at the grain size of the individual rather than the collective? What insights do we gain into the best ways to encourage productive reflection when we throw student characteristics into the mix? The next chapter will explore some of the effects of individual characteristics.