Microdevelopment

Transition Processes in Development and Learning

edited by

Nira Granott and Jim Parziale

CAMBRIDGE
UNIVERSITY PRESS

2002

10 Notebooks as windows on learning: The case of a science-into-ESL program

Rochel Gelman, Laura Romo, and Wendy S. Francis

Introduction

Our efforts to study the acquisition of new knowledge, especially in the areas of mathematics and science, take advantage of a combination of methods. These include experiments, controlled training studies, and qualitative analyses of videotaped classroom learning. In addition, we have been developing ways to treat the classroom as an in situ learning laboratory. Here we focus on a particular example, a science-into-ESL (English as a Second Language) classroom that functioned as a concept and language learning laboratory.

A key research goal was to gather "on-line" data on whether students moved on to and along relevant learning paths during the course of a semester. Given the classroom setting, we wanted these to be inconspicuous "probes" rather than more traditional learning test trials. Our solution shares important aspects of the microgenetic method, the topic of this book. We requested that students keep notebooks and make regular entries that could function as learning experiences for them and probes as to whether learning was happening over time. Notebook entries are a common feature of how scientists work, especially when it comes to trying to keep track of experiments and evolving ideas. They also have served historians of science as documents for analyzing the conditions of conceptual change. As we shall see, notebooks have the potential to serve researchers' efforts to obtain on-line information about students' progress in a non-intrusive manner. Thus, they can function both as a

An early report of the Science-into-ESL project was presented in a master's thesis by Romo (1993). The work covered here was supported by several funding sources, including grants from the Linguistic Minority Research Institute, University of California; the US Office of Education; and an NSF LIS grant entitled "Learning in Complex Environments by Natural and Artificial Systems." Correspondence about the chapter can be addressed to Rochel Gelman.

This chapter is dedicated to Anne Brown. She and her husband Joe Campione were ever-present conversationalists and cheerleaders from the very outset of the science-into-ESL project.

learning opportunity for students and as a research tool for investigators of concept and language acquisition.

Our commitment to a constructivist theory of concept acquisition motivated us to achieve on-line peeks at students' accomplishments (or lack of accomplishments) as their term progressed. Since learners are active participants in their own conceptual development, we cannot assume that they and their instructors will share a common interpretation of the lessons and related learning exercises. The knowledge and language skills they bring to a given subject influence what they attend to and how they interpret offered data. When the to-be-learned material is not readily related to what is already known, there is a risk that the inputs are ignored or even misinterpreted (Bartlett, 1932). Hence, it is critical to have ways to peek regularly at what, if anything, about the curriculum is being learned.

The science-into-ESL program featured here was developed by a team of high-school ESL and science teachers in collaboration with members of our university research team. A major goal of the program was to move ninth-grade students (fourteen and a half to sixteen years of age; median fifteen years) far enough along in their use and understanding of science in English so that they could take regular tenth-grade science classes, even though they had not completed their ESL course sequence. Of course, the program also had to increase students' overall English proficiency. At the time of study, ESL students at high schools in the Los Angeles Unified School District (LAUSD) were required to finish at least four semesters' worth of ESL instruction before starting science classes. This had the effect of limiting their opportunity to take enough years of high-school science courses to meet admission requirements to many colleges and other advanced levels of training.

The science-into-ESL curriculum used many of the lessons of cognitive science (see Bransford, Brown, & Cocking, 1999; Gelman & Gattis, 1995, for reviews). Every effort was made to ensure that students were active participants in their own learning, and that they engaged in guided discovery (Brown & Campione, 1996). In addition to being able to read newly prepared reading text materials about science topics, the students had opportunities to do collaborative experiments, make entries in their science notebooks, interpret data, talk and write about science, use computers, and interpret the course materials. Further, a group of bilingual UCLA undergraduates were present during voluntary after-school periods. Not only did the notebook requirement allow us to build in non-evaluative exercises that we could use to monitor students' learning over time, but it also fit with one of the desiderata of the course, this being that the students be able to do science and communicate about what was covered in readings and classroom lectures.

Monitoring learning in a classroom, given a constructivist perspective

Learners, be they infants, children, or adults, are not passive recipients of their environments. We all contribute to our own conceptual development by virtue of the fact that we use existing knowledge structures, no matter how skeletal they might be. Mental structures both require and contribute to epigenetic interactions with their environment. Without conscious effort, we more readily notice and remember those aspects of the environment that are related to available mental structures. Our ever-present tendency to assimilate data that are related to existing conceptual structures is often a good thing. Domain-relevant paths lead learners to inputs that are consistent with the structure of the domain to be learned; domain-irrelevant paths do not. Thus, existing structures of mind help novices move on to learning paths that are domain-relevant rather than domain-irrelevant.

There is a potential downside to the constructivist nature of our minds, namely, that it is always possible that beginning and novice learners will ignore and/or misinterpret their environments. Where mental structures have to be acquired de novo – as is surely the case for topics such as Newtonian mechanics, the theory of evolution, the stock market, etc. – learners have to acquire domain-relevant structures in addition to content (see Brown, 1990). Efforts to provide instruction in these domains must recognize and overcome a crucial challenge: learners may assimilate inputs to existing conceptual structures even when those inputs are inconsistent with existing structures and intended to force accommodation and conceptual change (Slotta, Chi, & Joram, 1995; Gelman, 1993, 1994). For this reason, students are often at risk of failure to learn mathematics or science. If we leave them to their own devices, they are biased to assimilate classroom inputs about fractions to their counting-based number theory. For example, they are likely to assume that 1/56 is less than 1/75 because 56 is less than 75.The risk of a mismatch between our intended inputs and what the learners take on is much increased when what is to be learned does not share structure with what is already known (Gelman, 1994; Gelman & Williams, 1998).

If the to-be-learned data do not map to an existing structure, the conditions are conducive for them to be regarded as irrelevant and therefore ignored (Stigler & Fernandez, 1995), or misinterpreted as relevant to an existing knowledge structure (Hartnett & Gelman, 1998). In the former case, learners fail to move on to the target learning path. In the latter case, they might even move on to the wrong learning path, one that encourages the development of misconceptions. These considerations highlight the need to have tools to determine whether learners move on to relevant learning paths when they are offered inputs meant to nurture new learning.

There is growing evidence that microgenetic studies – ones that probe for the kinds of hypotheses, language, and strategies children use at different points in the course of acquisition – contribute to efforts to understand concept development (e.g., Karmiloff-Smith & Inhelder, 1974/75, as well as the chapters by Goldin-Meadow & Alibali, Granott, Kuhn, Siegler, and Parziale in this volume). They are especially suited to gaining insight about the initial and changing nature of children's interpretations of inputs for learning. We had this advantage in mind when we set the goal of developing a version of such microgenetic analyses to chart learning paths in a school setting over the course of a semester. To do so we had to deal with three challenges.

The science-into-ESL materials were likely to be rather novel for many in the class. Therefore, despite extensive efforts to achieve a superb curriculum, it was not reasonable to expect the students to achieve learning with full understanding in a ten-week term. Mastery of the scientific concepts taught in high school can take many years. Acquisition of what Gelman and Williams (1998) dub "non-core domains" requires the assembly of new conceptual structures as well as the domain-relevant knowledge, and lengthy formal instruction is often required (see for example Bransford et al., 1999; Carey, 1991; Chi, 1992; Chi, Glaser, & Farr, 1988; Kuhn, 1970). We knew it would be foolhardy to expect the students to achieve learning with understanding. Therefore, we needed a design plan that differed from those microgenetic studies that track the evolution of strategy preference and deployment as learners converge on an endpoint, a correct solution, or a target learning plateau. The challenge then was to have ways to monitor whether students were at least attending to, and beginning to learn about, their lessons. A second challenge was to find a way to deal with the fact that, compared to most microgenetic studies, the time frame of our study was much expanded, indeed, to a full academic term. Finally, we knew that we could not achieve the level of experimental control that is a key feature of microgenetic studies. Still, our variant of the microgenetic method had to allow us to monitor whether or not students were moving on to and along relevant learning paths. Our search then was for a method that would provide a window on whether students were traversing learning paths which led toward science-relevant knowledge and language. As indicated, we responded to these challenges by taking advantage of a central feature of the science-into-ESL curriculum we were involved with – the students kept scientific notebooks. This was part of the plan to engage students in doing, talking, and reporting science.

Writing: A window on conceptual and language learning paths

The acquisition of scientific language and scientific concepts constitutes two sides of a coin; each is integrally dependent on the other (Kuhn, 1970; Halliday & Matin, 1993). When it comes to learning about technical domains, it is not stretching matters too far to treat science language learning as a form of second language learning. It is not just that the languages of science and mathematics are filled with specialized terms that derive their meaning from the domain within which they operate. There is also the problem that some everyday terms have completely different meanings. For example, the word "conductor" refers to the leader of an orchestra when used in everyday English. However, in the language of science "conductors" are substances or bodies that have the capacity to transmit or carry heat, electricity, etc. Within mathematics terms like "multiply" and "add" are more like false cognates; they do not have the same meaning in mathematics as they do in the natural language. They do not always mean "increase" or "get bigger," as shown in the result of multiplying two fractions or adding two negative numbers. Things are equally tricky when we consider the meaning of "add" when liquids are being combined. When one adds two cups of water to three cups of water, the result is five cups, but when one adds a 10 °C cup of water to another 10 °C cup, the resulting temperature is 10 °C. These examples illustrate that terms cannot always be treated as equivalent across different domains. Bransford et al. (1999) offer many further examples of the fact that learning about concepts and the specialized language (and notational system) of a domain go hand in hand. Such considerations reinforced our view that we would be able to probe whether students were moving on to and along a learning path that was related to the learning activities, by looking at their developing use of the domain-relevant scientific language.

Our decision to encourage students to write and ask questions about what they had been taught was also motivated by the finding that the study of explanations has provided researchers with a window into transitional states of students' mathematical and scientific understanding (Gelman, 1991; Hartnett & Gelman, 1998; Siegler & Jenkins, 1989). In addition, as individuals acquire knowledge about a topic, their ability to communicate and write about that topic progresses. For example, Siegler & Jenkins (1989) found that as children discover new strategies for addition problems their explanations are less articulate, that is, they use incomplete sentences, pause more often while talking, and engage in multiple starts and stops. These are taken to be indications that they are thinking about new and old strategies simultaneously. Related findings are provided by

Goldin-Meadow, Alibali, Church, & Breckinridge (1993). When children are in a transitional state of learning, the hand gestures that they use to describe their understanding of a math problem are often incongruent with their speech. This is because one kind of math strategy is being conveyed nonverbally while another competing strategy is conveyed through their verbal communication (Goldin-Meadow et al., 1993). Gelman and her colleagues have also illustrated the power of having students write about their understanding of to-be-learned concepts at different points in the learning process. Fourth- and fifth-grade children (nine- to eleven-year-olds) who were at different ability levels provided qualitatively different written answers when asked "why are there two numbers in a fraction?"

The foregoing examples illustrate another important point, this being that both mathematical understanding and language use differ as a function of how much students know about math. The same is true for learning about science. For example, Anderson and Shifrin (1980) demonstrated that when students acquire organized knowledge about spiders, they both understand and recall the salient features of a passage after reading it. Those who do not know about spiders recall poorly and have little understanding of the text.

Bransford et al. (1999) also review the growing evidence that teachers can benefit from clues about students' evolving and changing interpretations and understandings of the material being taught. An example of this comes from Stigler's analyses of video-taped lessons of elementary school mathematics teachers in Japan. The teachers made extensive use of students' changing explanations within and across lessons as a way to monitor students' progress. Stigler and his colleagues' characterization of the students' evolving solutions to problems brought out the fact that the students in the class often started out with different kinds of answers that covered a variety of areas. Over the course of classroom discussions, the teacher explored students' errors with them. Teachers then used these errors to guide students towards correct solutions and, from our point of view, move them on to relevant learning paths. (Stigler & Hiebert, 1999).

Aims and setting of the present study

As indicated, the curriculum was designed to embed science instruction into a ninth-grade ESL class in a public high school in the Los Angeles area. The high school which served the students in this study accommodates approximately 3,000 students, 20% of whom are designated as "Limited English Proficient." The school is in relatively good condition and has good sports fields as well as outdoor spaces. The great majority of the students in the school's ESL program speak Spanish, having come primarily from Mexico and El Salvador. Korean-speaking students comprise the next largest group; however, the program also serves students who speak Russian, Persian, Armenian, Tagalog, Vietnamese, Hebrew, Thai, Cantonese, and Mandarin. Many of the students traveled long distances from their homes on LAUSD school buses that picked them up at their neighborhood school. The data presented in this report were collected in an intermediate-level ESL2 class during the first semester in which the Science-into-ESL program was implemented.

Every effort was made to take into account the emerging views on how to apply the findings about the nature of conceptual learning to the design of classroom instruction. An effective program for preparing students with the relevant background for learning new science concepts requires more than just having them memorize terms or lists of facts (e.g., Beck & McKeown, 1989; Carey, 1988; Dooling & Lachman, 1971; Glynn, Yeany, & Britton, 1991). The ability to repeat the definitions of terms does not guarantee a correct understanding of the concepts (e.g., Osborne & Cosgrove, 1983; Stepans, Dyche, & Beiswenger, 1988). Therefore, it is important to attempt to relate classroom reading and other instruction materials and tools around a plan to develop into class an organized knowledge base of related concepts. It was agreed that the program would do well to concentrate on a limited number of topics in some depth. Topics from the biological and physical sciences that met school system guidelines, including those in the *California State Board of Education Science Framework* (State of California, 1990), were organized into a ten-unit program that focused on the conceptual themes of variability, energy, interdependence, and change. These were featured in the reading materials prepared for the classes (Meck & Gelman, 1992).

Materials were ordered to have latter lessons build on earlier ones. When it seemed warranted, several units were devoted to a topic. The titles of the initial ten units were: 1. Sun; 2. Photosynthesis; 3. Respiration; 4. Local Winds; 5. Temperature and State; 6. Buoyancy and Density; 7. Water Cycle; 8. Food Energy; 9. Organs and Organisms; 10. Interactions and Ecosystems. Each unit consisted of:

i. an *initial reading* designed to focus the student on the core concept(s) of the unit and to provide the linguistic means to respond appropriately to a short pretest for that unit
ii. a short *pretest* of a representative core concept presented in the initial reading; for the teachers, a *listing of the core concepts and vocabulary* around which each of the units was built

iii. a *main reading* that incorporated the core concepts and vocabulary into a text. This was paired with reading comprehension and second language development exercises
iv. a *laboratory exercise* focusing on the core concepts and including both demonstrations and experiments that the students did in groups, using inexpensive materials brought from home or a grocery. For each laboratory exercise, there was a teacher-oriented set of instructions
v. a *review* involving language and concept development exercises
vi. a short *follow-up test* to complement item ii, the short pretest
vii. a *journal notebook* for students to keep track of their classwork (see below)

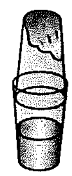
Given the findings that opportunities to do science improve the likelihood of learning science (Brown & Campione, 1996), we created groups of "experimenters." Students were assigned to groups for conducting experiments, talking about them, and recording their observations. These groups also prepared their findings to report to the class, and the teacher was encouraged to engage students in guided question-answering opportunities. Also, as indicated, notebooks were an integral part of the students' participation in the class. They were used to make notes about a variety of learning experiences, including the preparation of graphs and other scientific figures, record keeping and analysis of experimental data, lessons on science vocabulary, reviews, and question-asking. Two examples of the kinds of notebook entries students made to record their experiments are show in a schematized version in figure 10.1. Figure 10.2 illustrates the kind of writing activities that were part of the review sessions that ended a unit.

The review session for a unit started with a group discussion during which the group generated a list of ten to twenty science terms relevant to the unit's theme. After the class agreed on the list of relevant terms, students were given approximately fifteen minutes to construct concept maps in their notebooks. In this exercise, students drew lines to link concepts that they thought to be related to each other. Subsequently, they were given fifteen minutes more to write up to ten sentences to describe how they had inter-related the concepts in their concept maps. The sentences that students did write became the focus of our analyses of whether students' knowledge of the concepts and language of science had indeed moved on to a learning path by the end of the second half of the course.

We turn to an extended consideration of the kinds of analyses we did with the review sentences. We also present some standard outcome measures. However, evaluation of the curriculum itself is not the main focus

Student A November :20: 199–

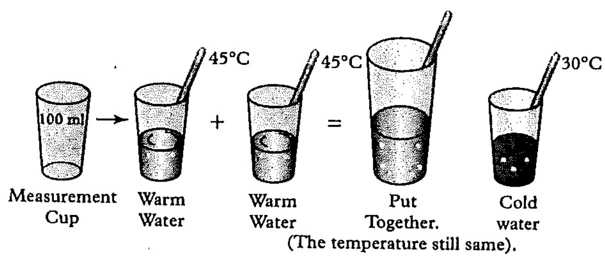
Illustration about the Experiment



→ Clear cup became have water vapor from heat water.

→ White cup contents hot water became to heat the other cup

Water + heat = Gas.
Gas – heat = liquid.



Student B October 14, 19–

<u>Conclusion</u>

1. The water absorbed more light in the sun.
2. Covered glasses generally got warmer.
3. Covered glasses absorbed more heat than covered.
4. Black absorbs more than white.

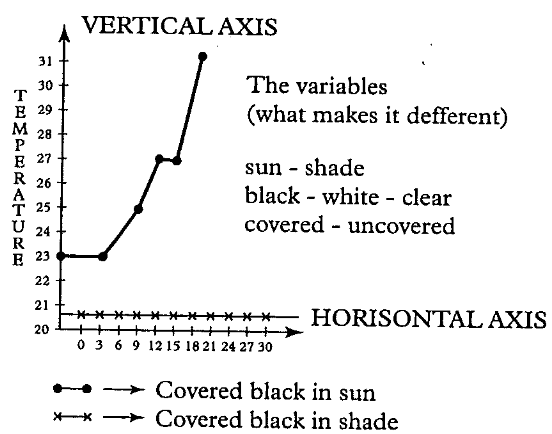
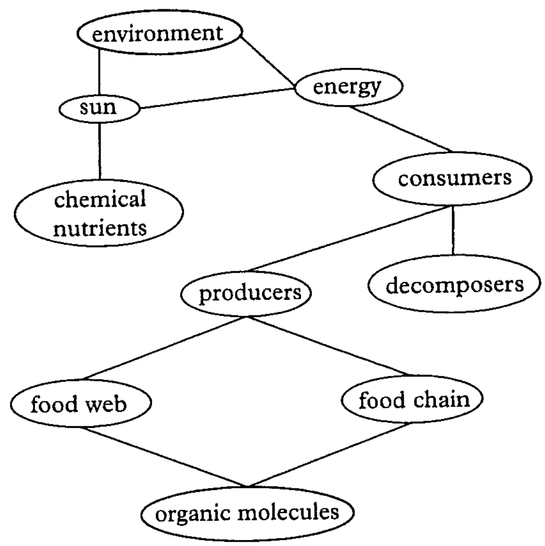


Figure 10.1 Redrawn examples of the kind of notebook records that Students *A* and *B* made for the experiments they did to accompany their reading. (Students' spelling and grammatical errors have not been corrected.)

Student C

Interaction
Cognitive map
6/2/9-



1) Most of the energy comes from the sun.
2) The sun is necessary for the environment.
3) Descomposers get the energy that they need from organic molecules.
4) Produsers need the sun energy to survive.
5) Produsers are called food webs.
6) Chemical nutrients are bad for the environment.
7) Descomposers need organic molecules to survive.
8) Consumers need chemical nutrients.
9) Not all the energy come from the sun.
10) Consumers need food chain.

Figure 10.2 A redrawn example of a notebook entry for a given review session. (Students' spelling and grammatical errors have not been corrected.)

in this chapter. It is instead to show that investigators can use students' notebooks to accomplish some of the key advantages of the microgenetic method, namely, that notebooks are windows into which to regularly peek at "on-line" learning (or non-learning) over an extended period of time and in a regular classroom setting. The kinds of on-line measurements that we have been developing need not be tied to a particular instantiation of a programmatic effort to embed relevant science learning materials into an ESL program.

Analyzing the progression of students' ideas: The written sentences

A total of thirty students were enrolled for at least some part of the term. Of the thirty students, eight either moved, started the semester late, or were excessively absent. The notebooks of the remaining twenty-two students who were present throughout the semester provided the corpora for our analyses.

The writing exercises for Units 1 and 10 were excluded from analysis because the first unit served as practice and the last unit was separated from the other nine units by a holiday break. The remaining eight units yielded a database of 1,271 sentences. Of these sentences, thirty-seven (all produced by the same student) were judged to be inappropriate, in that they stated simply which terms in the map could be connected, without describing the relationship. Two independent coders evaluated whether or not each of the remaining 1,234 sentences was comprehensible and what the intended message was. The coders' classifications matched in 98% of the cases (Cohen's $k = 0.89$). Consensus was reached on all discrepancies, resulting in the exclusion of eighty-one sentences (6.6%) from further coding, whereupon the remaining body of codeable data contained 1,153 sentences.

Some students did not complete the task of writing ten sentences in the time allotted. Students who consistently wrote at least five sentences in every exercise were assigned to a High-Production group. The other students were assigned to a Low-Production group, the rationale being that students less fluent in English might perform differently owing to the extra mental effort it requires for them to produce language (Francis, 1999). Half of the twenty-two students generated at least five sentences during each review; the other half of the students consistently wrote fewer than five sentences in every exercise. Because this difference in production might lead to different levels of overall performance, it was used as a blocking variable in the statistical analyses. For the purposes of analysis, the groups that produced more than five sentences per exercise were called the "High-Production Group" and those who produced fewer were called the "Low-Production Group." As it turned out, the levels of performance of these two groups did not differ for most variables and effects analyzed. Therefore results for the High- and Low-Production students are reported separately only when differences or interactions were present.

Figure 10.3 contains examples of individual students' tendencies to produce multi-clause sentences as a function of successive curriculum units. Although the shapes of the individual learning curves vary across students, overall students tended to generate sentences that were more

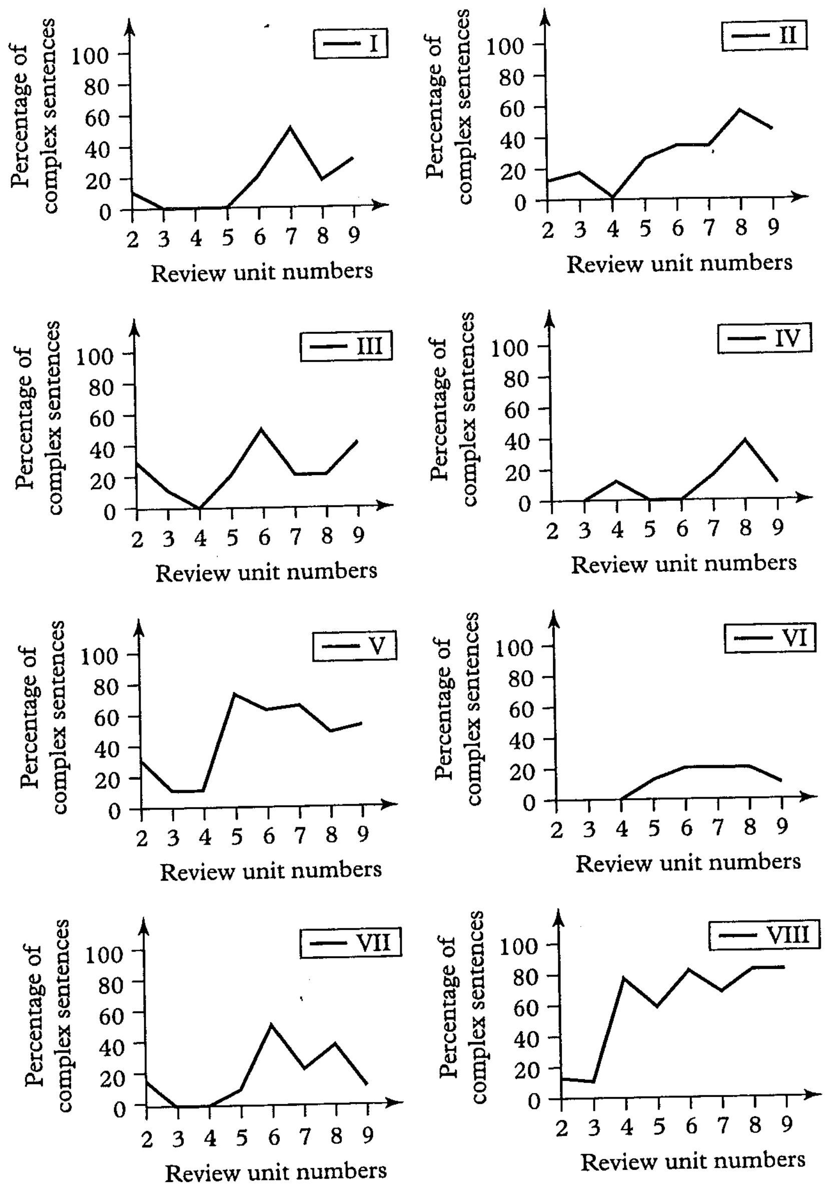


Figure 10.3 Eight individuals' learning curves for percentage of complex sentences as a function of review unit. Unit 1 is not included because it was used to introduce the exercise.

structurally complex in the latter half of the science course. The tendency to move forward was our first clue that the review sentences possibly contained more specific science information as a function of experience. We will return to the matter of the possibility of different acquisition functions in the discussion. Now we focus on ways to measure whether the students

Table 10.1 *Concept specificity coding system and examples*

Level	Content specificity category	Examples of student sentence productions
1	Properties or category membership	The sun is a star. Sunlight has different colors. Stars have their own heat. Vapor is a type of gas.
2	Needs or sources	Energy comes from the sun. All plants need energy. Our body needs a lot of carbohydrates.
3	Mechanics or functions	The sun gives energy to the plants. Animals breathe oxygen. Air is always pulled by gravity. Green plants make food from the sunlight.
4	Conditions or purposes	A hot air balloon is sometimes dangerous, because it has gas in it. All plants need energy in order to survive. Hot air goes up, because it is less dense. When we add heat energy to the water, it causes the water to change temperature.

were on relevant learning paths at least by the second half of the term. To accomplish this goal we designed a coding system to quantify both (1) the syntactic complexity of the sentences and (2) the amount of conceptual information expressed in the sentences over time (see table 10.1).

Analysis of sentence syntactic structure

The sentence structure analyses focused on two variables relevant to writing sentences with complex ideas: 1) the use of embedded clause structures and 2) the inclusion of prepositional phrases. The first kind of complexity in sentence structures is common in scientific language, such as "if you boil water, it vaporizes" (Celce-Murcia & Larsen-Freeman, 1983). Therefore to describe specific concepts, the students would need to develop their skills in writing these kinds of sentences. The addition of prepositional phrases serves to expand the number of concepts expressed within a sentence.

Each sentence was coded as either *simple* or *complex*. A *simple* sentence was made up of a single main clause (e.g., "All plants need energy"). A *complex* sentence was made up of at least one main clause and one embedded clause (e.g., "All plants need energy in order to survive"). Sentences joined together simply by coordinating conjunctions (e.g., "and," "or") were counted and coded as separate sentences. Each sentence was coded for whether or not it contained any prepositional phrases. For each student, the proportion of complex and simple sentences and the proportion of sentences containing at least one prepositional phrase was calculated for each half of the course.

The total number of sentences produced in the first and second halves of the course were 555 and 598, respectively. The mean proportion of complex sentences increased from 13% in the first half to 31% in the second half ($F(1, 19) = 38.068$, $MSE = 0.009$, $p < 0.001$). The average proportion of sentences containing one or more prepositional phrases also increased from 34% in the first half to 48% in the second half of the course ($F(1, 19) = 14.95$, $MSE = 0.01$, $p < 0.001$). These results indicate that the sentences in the latter half of the course were made up of complex phrase structures, a condition in English for writing more conceptually informative prose. The next set of analyses addressed whether the syntactic complexity of the sentences served this purpose.

Analysis of conceptual information

Students described and explained how the science concepts/terms related to one another in a variety of ways across sentences. In the analysis, four categories were identified (see table 10.1). At the most general level, *Properties or Category Membership* (Level 1) sentences described attributes associated with at least one entity. Next, *Needs or Sources* (Level 2) sentences described a simple relationship of two or more entities needing one another, or one deriving from the other (e.g., energy sources or food sources). *Mechanics or Functions* (Level 3) sentences provided more information by describing how entities act or operate on other entities, but no reasons or conditions were specified. The most conceptually informative sentences, *Conditions or Purposes* (Level 4) sentences, explained conditions and circumstances for how and why entities, substances, or processes were related to one another. Table 10.1 shows examples of sentences that students produced in each of these four categories. Two independent raters coded the conceptual information of each sentence, which ranged from general (Level 1) to more conceptually informative (Level 4). Agreement on classifications into the four categories was 94% (Cohen's $k = 0.92$). Consensus was reached on all discrepancies.

Table 10.2 shows the distribution of sentences across the four category levels. In the first half of the course (Units 2 through 5), the most prevalent kind of sentence was a *Properties or Category Membership* sentence ($M = 47\%$). It seems that during the first half of the course students mostly attended to properties of entities and wrote general definitions.

Table 10.2 *Average proportion (standard deviation) of sentences at each concept specificity level*

Level	Content specificity category	1st half of course	2nd half of course
1	Properties or category membership	0.47 (0.14)	0.26 (0.13)
2	Needs or sources	0.15 (0.14)	0.12 (0.10)
3	Mechanics or functions	0.20 (0.09)	0.25 (0.11)
4	Conditions or purposes	0.18 (0.13)	0.37 (0.18)

In the second half of the course (Units 6 through 9), *Conditions or Purposes* sentences were most prevalent ($M = 37\%$). Students shifted their focus from describing general attributes of entities to describing specific conditions, reasons, or circumstances governing relationships among entities.

To further analyze this pattern, an *average concept specificity* measure was computed for each student for each half of the course. This measure was derived by summing products of category levels and their corresponding proportions.[1] This weighted average provided a useful way to compare performance across the two halves of the course (though it clearly has only ordinal, not interval scaling properties). Average concept specificity scores increased from 2.08 in the first half to 2.72 in the second half [$F(1, 19) = 113.92$, $MSE = 0.033$, $p < 0.001$], indicating that students expressed more specific ideas about the science terms in the second half of the course. This change reflects an increase in the production of *Conditions or Purposes* sentences from 18% of all sentences in the first half to 37% of all sentences in the second half. All but one student produced several examples of this type of sentence. Thus, it appeared that as students became familiar with the properties and functions of entities, they shifted to writing more sentences about why and for what purpose these entities stood in relationship to one another. It is important to note that although the proportion of attribution sentences decreased from the first half to the second half, the students continued to write some definitional sentences, perhaps because they were introduced to novel concepts in new units or were still clarifying the older ones.

What kinds of conditions, reasons, or circumstances were described in these *Conditions or Purposes* sentences? Again, because this was an English course and not a regular science course, we did not expect the ideas to be extremely sophisticated. Overall, the students wrote mostly

[1] For example, for a student who produced 15% Level 1, 20% Level 2, 30% Level 3, and 35% Level 4 sentences, the weighted average would be calculated as follows: $(1 \times 0.15) + (2 \times 0.20) + (3 \times 0.30) + (4 \times 0.35) = 2.85$.

about conditions or reasons (Category Level 4) associated with mechanics or functions of things (60% of 336 sentences). For example, students explained that "When we add heat energy to the water, it causes the water to change temperature." Nineteen percent of the sentences provided reasons or conditions for needs and sources (e.g., "Plants need water, carbon dioxide, and sunlight for photosynthesis"). Twenty-one percent of the sentences described conditions or circumstances for a property or a category membership (e.g., "The balloon has gas when it has air"). Agreement on these subcategories was 92% (Cohen's $k = 0.87$).

Accuracy of the content

To measure the accuracy of the sentence content, each sentence was scored on a scale from 0 to 4, with higher numbers indicating a more accurate display of conceptual understanding. Scale anchor points were 0 = no display of understanding; 1 = mostly unclear and inaccurate; 2 = about half correct and half incorrect; 3 = mostly clear and accurate; 4 = clearly accurate display of understanding. Two coders, who had assisted in the classroom, worked together to score the accuracy of each sentence. They agreed on 90% of the cases.

For each student, an average accuracy score was computed for each half of the course by dividing the total score across the units by the total number of sentences (i.e., weighting each sentence equally). On average, students' sentences were scored as mostly clear and accurate, a score of approximately 3. The production groups differed in their patterns of content accuracy across the two halves of the course (F (1,19) = 14.21, $MSE = 0.03$, $p < 0.001$). Accuracy of the Low-Production Group did not change significantly from the first to the second half of the course (from $M = 2.93$ to $M = 2.83$); but they did maintain their level of accuracy. In contrast, the High-Production Group improved significantly in accuracy from the first to the second half of the course (from $M = 2.90$ *to* $M = 3.20$, $p < 0.005$), suggesting that students who could produce more sentences showed higher levels of science accuracy.

Ability to explain new phenomena: The unit quizzes

Although the content of students' sentences was mostly accurate, one could argue that students could have been simply repeating facts that they had heard in class. To rule out this alternative explanation for their high level of content accuracy, we compared the students' scores to another source of written data – their explanatory answers to the posttest quizzes that accompanied each unit. The quizzes were designed, as part of the curriculum, to measure whether students could transfer their knowledge of key concepts of the units to explain novel scientific phenomena. The questions and the target answers had not been reviewed in the curriculum, although an analogous situation might have been discussed. Thus, these tests provided us with a written measure of the students' understanding of the science material that could not have been simply copied or repeated from class materials or discussions. For example, following the unit on buoyancy and density, the students read the following passage and were asked to write a few sentences to explain the various phenomena described.

> Hot air balloons are dangerous but very exciting. Sometimes, we can see only one balloon high in the air, and sometimes we can see many balloons together. The people who ride a balloon stand or sit in a large basket below the balloon. As they move through the air, pushed by the winds, they fall gently and slowly toward the ground. When the people in the basket want to go up, they turn on a burner below the balloon and for a short time shoot a hot flame up into it. The balloon slowly climbs higher. After a while, the balloon again goes slowly toward the ground. How does a hot air balloon work? The mouth of the balloon is open to the air – there is nothing inside the balloon except air. What causes the balloon to go up?

Students' responses to these questions were scored independently by the ESL instructor and by a teaching assistant, who agreed on 94.5% of their ratings. Two points were given for responses in which the science information was clearly accurate with respect to the question; one point was given if most but not all of the science information was accurate; and zero points were given if the science information was clearly inaccurate.

For each student, a percentage score was computed across all completed unit posttest quizzes by dividing the sum of their unit quiz scores by the number of unit quizzes completed ($M = 33\%$, $SD = 22\%$). All of the students completed at least seven quizzes. We reasoned that if students were learning more about science concepts, and they could describe their understanding accurately in their sentences, they should also be able to provide coherent, accurate responses to questions on the unit quizzes. Indeed, the quiz scores and the content accuracy scores were highly correlated ($r = 0.79$, $p < 0.001$). This result provides some independent evidence against the explanation that the accurate concepts expressed in students' sentences were merely copied or repeated without understanding.

Requests for more science information

At the end of each unit, the instructor asked students to write down any questions they had about the material and what they would like to learn

Table 10.3 *Examples of specific "why" and "how" questions asked by students*[a] *(corrected for spelling and simple grammar)*

- why we can't drink the water in the ocean because when you put water in a pot and put salt, we can drink the water
- why people can't live underwater and why other animals can't live on the land
- why the dirt gets cold more faster than water
- how the winds go around
- why if we eat a lot of calories, we will get fat
- how come hot air goes up and changes to cold air
- how a tree can live in the winter
- why any object with density that is greater than water will sink and any object whose density is less than water will float
- how do we know that green plants have covered our planet for over a billion years?
- why can't we breathe in the water?
- why is the sun very hot?
- if gravity pulls everything's down including the air, why doesn't it kill us?
- how can a tree pick up water from his root?
- how do living things get energy from food?
- I would like to find out if there is someone in space like ET.
- what would happen if a space man takes off their clothes in space?

[a] Some students failed to ask questions, despite the instructions for this exercise. Some examples of what these students wrote are included here.

about in the future. All of the students provided at least one request for additional specific information about a scientific process or phenomenon (e.g., "Why are the hot air and cool air different weights?" or "How do the stars make their own light?") over the course. Most of the questions were about why or how things function, take place, or come to be (72%). Some examples of the students' questions (corrected for spelling and simple grammar) are presented in table 10.3. The remaining responses included general statements either restating the title of the unit (e.g., "I don't understand about local winds"), or general comments about wanting to learn more science (e.g., "I would like to learn more about respiration"). Two independent raters coded the number of specific information requests for each of the students (Cohen's $k = 0.98$ for each question). The mean number of information requests for each of the twenty-one students (missing data on one student) was computed across all completed units, with a maximum possible score of 18 ($M = 9.52$, $SD = 5.01$).

A correlational analysis was performed to examine the relationship of the number of specific information requests to the concept specificity and correctness scores from the sentences. Students who asked more specific kinds of questions also included more correct information in their sentences ($r = 0.46$, $p < 0.05$), although asking specific questions was not related to sentence concept specificity. Furthermore, increased specific question-asking was reliably associated with students having more correct science explanations in the unit quizzes ($r = 0.56$, $p < 0.01$). Overall, it seems that students who gained a better understanding of the material were also able to ask more thoughtful, focused questions about the science topics throughout the course.

Question-asking may also be considered an indicator of students' interest in science or motivation to learn science. Indeed, there was a marginally positive correlation between how many specific questions the students asked and their ratings on a seven-point scale ($M = 5.6$, $SD = 1.0$) of how interested they were in learning science ($r = 0.42$, $p = 0.057$). Although the number of specific questions did not predict whether or not they took a mainstream science class the following year, there was interesting relation to the grades of the twelve students who did go on to take science. Those students who had asked more questions about the science material in our program the previous year did better in their regular science courses the next year ($r = 0.62$, $p < 0.05$). The mechanism for this relationship could be motivation, or it could be that the students continued to ask specific questions in their regular science classes to clarify their understanding.

English grammar skills

A main concern among the ESL instructors at the school was that the new science curriculum would use up valuable class time that would otherwise be dedicated to basic grammar drills and exercises and that their students would fall behind in learning English. Given our plans for coding the sentences, we gained information that addresses these concerns (Francis, Romo, & Gelman, in press). A brief discussion of the English language-learning results is included in this section.

To start, we simply scored whether each of the 1,153 sentences contained a grammatical error or not. The number of sentences in each half of the course that were free of grammatical errors was divided by the total number of sentences in each half to obtain the proportion of sentences that were completely grammatical in each half. The percentage of error-free sentences decreased from 35% in the first half to 24% in the second half of the course (F (1, 19) = 10.03, $MSE = 0.01$, $p < 0.01$). As expected, the ESL students had problems producing sentences with correct morphosyntax throughout the semester. In

Table 10.4 *Correlations among the five main concept and language variables across students*

	Content accuracy	Clause complexity	Preposition frequency	Whole-sentence grammaticality
Content specificity	0.73[c]	0.80[c]	0.62[b]	0.07
Content accuracy		0.78[c]	0.52[a]	0.24
Clause complexity			0.49[a]	0.35
Preposition frequency				0.24

[a] $p < 0.05$; [b] $p < 0.005$; [c] $p < 0.001$

particular, students made errors in applying subject–verb agreement, auxiliary and modal verbs, prepositions, and determiners. The increase in sentences with grammatical errors was most likely due to the increase in opportunities to make more errors as the students wrote longer and more complex sentences. Surely many of us have extremely accomplished colleagues who do and write science in English but who nevertheless fail to include an obligatory determiner or mix up their use of "he" and "she," and so on. Although we expected the students to continue to have persistent problems producing correct morphosyntactic constructions, we thought it possible that these kinds of grammatical errors would not be related to students' learning of information about science.

Relationships among the conceptual and language variables

The results of the analyses reported in the previous sections motivated us to examine the relationships among the conceptual and linguistic variables. We selected five variables for a correlational analysis: *Content Specificity* (scaled from 1 to 4), *Content Accuracy* (scaled from 0 to 4), *Clause Structure or Complexity* (simple or complex), *Frequency of Prepositional Phrases* (none as opposed to one or more), and *Whole-Sentence Grammaticality* (correct or incorrect). These relationships were examined to reveal which features tend to go together to characterize individual students' ideas and written expression.

For the correlational analysis, we obtained the average score assigned to each of the five variables across all the sentences produced by each student. (For variables based on dichotomous classifications, the averages correspond to proportions.) Table 10.4 shows the correlations among the five variables. It can be seen that students who wrote about specific complex ideas also wrote more accurate ideas. Students who wrote more specific and/or accurate ideas also used complex sentence structures that contained more embedded clauses and more prepositional phrases. It appears that students generated sentence structures that were useful, and sometimes necessary, to explain the conditions, circumstances, reasons or purposes of scientific phenomena.

In contrast, whole-sentence grammaticality was not significantly associated with any of the content or sentence structure variables. It is important to note, however, that the directions of these nonsignificant correlations were positive – one cannot say that students who express specific complex or accurate ideas or who use more embedded clauses and prepositional phrases were compromised on their morphological marking. That is, there was no trade-off between learning science and learning English morphology, or between learning to use complex sentence structures and learning English morphology.

The fact that overall grammatical accuracy score is a poor predictor of one's ability to engage in learning about science-relevant concepts and communication skills deserves special attention. For all students in the ESL program, including those in ours, took the LAUSD's end-of-semester evaluation exam, which is a multiple-choice test on English grammar. Our students' scores were comparable to those of former ESL students who were taught the standard ESL curriculum by our teachers. Much to the relief of the ESL instructors, the students who participated in this program performed as well as had previous ESL2 students (Gelman, Meck, Romo, Meck, Francis, & Fritz, 1995). True, the ESL scores mean that learning about science did not come at the cost of learning English grammar as indexed by expected levels of performance on standardized ESL tests. However, if the only measure used to assess students' progress were these tests, then the students in our program would not have had an opportunity to start on to science learning paths. Given that 64% of the students in our study actually went on to take a tenth-grade science course that is taught in English with English textbooks, our analyses of different aspects of language illustrate the importance keeping in mind the fact that "language" is not a unidimensional domain.

Some final thoughts

There is an overarching lesson about the kinds of results presented here, one that dovetails well with the patterns of results presented in a number of chapters in this book. This is that the acquisition of knowledge is a multi-faceted matter, so much so that different measures taken at the same time even can diverge, conflict, or disagree. In our case, we show that students' ability to write about science-relevant concepts improved from the first half to the second half of the course. So did their ability

Figure 10.4 The learning curve that results when the individual data points in figure 10.3 are averaged.

to write with complex sentences of the kind that support the expression of if–then and causal scientific content. Still, this does not mean that the production of these sentences was perfectly correct at the syntactic level. Indeed, whole-sentence grammaticality was not correlated with any of our content or sentence structure measures. This means that, for example, students who earned relatively high scores on measures of conceptual complexity still could have made errors regarding determiners, plural markers, and even subject–verb agreement.

As indicated, this chapter has focused on the idea that students' science notebooks provide a rich data source for monitoring what, if anything, students are starting to learn about. The analyses we presented illustrate the feasibility of this approach; however, these are by no means the only analyses that can be done. We chose to focus on first-half/second-half analyses. Some will ask why we did not do group, trial-by-trial analyses. An example of such an analysis would be like that shown in figure 10.4, where we show the average learning curve for the eight individual curves shown in figure 10.3. The averaged data give the impression that use of complex sentences improved rather gradually, starting around Unit 4. However, if we compare figures 10.3 and 10.4, it becomes clear that this is not true for all of the individuals represented in figure 10.3. That is, it seems as if the different students had different acquisition curves. The average curve obscures this. It is because there is an obvious first-half/second-half difference across the individuals that we started at this level of comparison.

Now we have some ideas about the kinds of things that students were or were not learning to write about in the Science-into-ESL program. We expect that future analyses could help us get a closer picture of the possible relationships between the ability to use the kinds of syntactic structures that support the expression of scientific knowledge and the learning of the scientific knowledge itself. Indeed, we think it could be possible to focus on the students' entries for their science experiments. The examples in figure 10.1 are quite characteristics of the notebooks as a set. This encourages us to consider whether to provide the opportunity to track the students' use of graphs, notational systems, experimental materials, and interpretation of experiments.

We end with a word of caution. It is unlikely that any kind of notebooks can be used to monitor learning paths and their shapes. It is extremely important to think about the kinds of entries that students might make and how these facilitate the joint goals of moving students on to learning paths and having a way for teachers and experimenters to peek to see if this is happening.

References

Anderson, R. C., & Shifrin, Z. (1980). The meaning of words in context. In R. J. Spiro, B. C. Bruce, & W. F. Brewer, (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Erlbaum.

Bartlett, F. (1932). *Thinking: An experimental and social study*. New York: Basic Books.

Beck, I. L., & McKeown, M. G. (1989). Expository text for young readers: The issue of coherence. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 47–67). Hillsdale, NJ: Erlbaum.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (1999). *How people learn: Brain, mind, experience and school*. Washington, DC: National Academy of Sciences.

Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science, 14*, 107–133.

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Contributions of instructional innovation to understanding theory* (pp. 229–270). Hillsdale, NJ: Erlbaum.

Carey, S. (1988). Why John and Jane aren't learning science. In *Science and public policy seminars*. Federation of Behavioral, Psychological, and Cognitive Sciences. Washington, DC: American Psychological Association.

Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), *Epigenesis of mind: Studies in biology and cognition*. Hillsdale, NJ: Erlbaum.

Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The grammar book: An ESL/EFL teacher's course.* Rowley, MA: Newbury House.

Chi, M. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science.* Minneapolis, MN: University of Minnesota Press.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.) (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Dooling, D. J., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology, 88,* 216–222.

Francis, W. S. (1999). Analogical transfer across languages in English–Spanish bilinguals. *Journal of Memory and Language, 40,* 301–329.

Francis, W. S., Romo, L., & Gelman, R. (in press). Syntactic structure, grammatical accuracy, and content in second-language writing: An analysis of skill learning and on-line processing. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing.* n.p.: Elsevier.

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gelman (Eds.), *Epigenesis of mind: Essays on biology and cognition* (pp. 293–322). Hillsdale, NJ: Erlbaum.

Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. Medin (Ed.), *Learning and motivation, Vol. 30* (pp. 61–96). New York: Academic Press.

Gelman, R. (1994). Constructivism and supporting environments. In D. Tirosh (Ed.), *Implicit and explicit knowledge: An educational approach* (pp. 55–82). Norwood, NJ: Ablex.

Gelman, R., & Gattis, M. Lee (1995). *Trends in educational psychology in the United Statess. Recent trends and developments in educational psychology: Chinese and American perspectives.* Paris, France: UNESCO.

Gelman, R., Meck, G., Romo, L., Meck, B., Francis, W., & Fritz, C. O. (1995). Integrating science concepts into intermediate English as a second language (ESL) instruction. In R. F. Macías & R. G. García-Ramos (Eds.), *Changing schools for changing students: An anthology of research on language minorities, schools, and society* (pp. 181–203). Santa Barbara, CA: UC Linguistic Minority Research Institute.

Gelman, R., & Williams, E. (1998). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In W. Damon (Series Ed.), & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology, Vol. 2: Cognition, perception, and language* (5th edn., pp. 575–630). New York: Wiley.

Glynn, S. M., Yeany, R. H., & Britton, B. K. (1991). A constructive view of learning science. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), *The psychology of learning science* (pp. 3–19). Hillsdale, NJ: Erlbaum.

Goldin-Meadow, S., & Alibali, M. W. (this volume). Looking at the hands through time: A microgenetic perspective on learning and instruction.

Goldin-Meadow, S., Alibali, M. W., & Church, R. Breckinridge (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100,* 279–297.

Granott, N. (this volume). How microdevelopment creates macrodevelopment: Reiterated sequences, backward transitions, and the Zone of Current Development.

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power.* Pittsburgh, PA: University of Pittsburgh Press.

Hartnett, P. M., & Gelman, R. (1998). Early understandings of number: paths or barriers to the construction of new understandings? *Learning and Instruction: The Journal of the European Association for Research in Learning and Instruction, 8,* 341–374.

Karmiloff-Smith, A., & Inhelder, B. (1974/75). If you want to get ahead, get a theory. *Cognition, 3,* 195–212.

Kuhn, D. (this volume). A multi-component system that constructs knowledge: Insights from microgenetic study.

Kuhn, T. S. (1970). *The structure of scientific resolutions.* Chicago, IL: University of Chicago Press.

Meck, G. H., & Gelman, R. (1992). *Interrelated science concepts in English as a Second Language: A science content text for intermediate level ESL instruction.* Draft, Department of Psychology, UCLA.

Osborne, R. J., & Cosgrove, M. M. (1983). Children's conceptions of the changes of state of water. *Journal of Research in Science Teaching, 20,* 825–838.

Parziale, J. (this volume). Observing the dynamics of construction: Children building bridges and new ideas.

Romo, L. F. (1993). Science concept development and language proficiency in ESL instruction. Unpublished master's thesis. University of California, Los Angeles.

Siegler, R. S. (this volume). Microgenetic studies of self-explanation.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies.* Hillsdale, NJ: Erlbaum.

Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition & Instruction, 13,* 373–400.

State of California (1990). *California State Board of Education Science Framework.* Sacramento, CA: California Publication.

Stepans, J. I., Dyche, S., & Beiswenger, R. (1988). The effect of two instructional models in bringing about a conceptual change in the understanding of science concepts by prospective elementary teachers. *Science Education, 72,* 185–195.

Stigler, J. W., & Fernandez, C. (1995). Learning mathematics from classroom instruction: Cross-cultural and experimental perspectives. In C. A. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development. The Minnesota Symposia on Child Psychology, Vol. 28* (pp. 103–130). Mahwah, NJ: Erlbaum.

Stigler, J., & Hiebert, J. (1999). *The teaching gap.* New York: Simon & Schuster.