TOWARD LEARNING PROGRESSIONS AS TEACHER DEVELOPMENT TOOLS

While learning progressions (LPs) represent an important advance beyond standards-style lists of what students need to know in science, creation of the progressions alone may not be sufficient to realize change in teacher practice, or to help students learn. However, if LPs are accompanied by information that will scaffold teachers' development of the knowledge they need to teach a particular scientific concept, their impact on science education has the potential be amplified. To inform the design of LPs that will represent the knowledge necessary to effectively teach the big ideas in science, this paper presents four design criteria for LPs for teacher knowledge, suggests research frameworks for the development and validation of such LPs, and then uses the case of an ongoing research project to highlight challenges faced in such projects exploring how LPs might be used by practicing teachers. The struggles represented by the teachers in this research project present a microcosm of the need for the educational research community to focus on its 'user' community when developing learning progressions.

Erin Marie Furtak, University of Colorado at Boulder

Background

A good teacher is not just a teacher, and not just a science teacher, but a biology teacher, a chemistry teacher, or an astronomy teacher (Boz & Boz, 2008; Geddis, 1993). One may even argue that even a biology teacher, as a result of his or her training in the multiple disciplines of biology, might be better prepared to teach anatomy as compared to ecology. In fact, in order to successfully navigate students' ideas during instruction, teachers need to develop not only their subject-specific knowledge, but also knowledge about how students learn the subject (Gess-Newsome, 1999; Shulman, 1986).

The field of science education has created several documents in the last two decades that have attempted to catalog the immense content knowledge that students are expected to learn. The most influential of these documents are the National Research Council's *National Science Education Standards (NSES)* (National Research Council, 1996) were the first national attempt by the US to standardize the scientific content students should learn at every K-12 grade span. Modeled on the National Council of Teachers of Mathematics Standards (2000) and succeeding the American Association for the Advancement of Science's *Science for All Americans* and *Benchmarks for Science Literacy* ([AAAS],1990, 1993), the NSES were intended to standardize science learning across the US, and to provide a framework for teachers, schools, and districts to follow when planning for instruction.

However, according to Collins (1998), the NSES were not intended to be a how-to document, and as a result, they do not offer specific guidelines to teachers about what students should learn, or how. The NSES are grouped by grade level (K-5, 6-8, and 9-12) and content area (Life Science, Physical Science, etc.), and are a list of correct ideas that are neither listed in sequence of when they should be learned, nor are they based on development of student understanding.

The AAAS' *Atlas of Science Literacy* (2001) made a herculean attempt to reorganize information from *Benchmarks for Science Literacy* and the *NSES* according to conceptual themes rather than grade spans; however, the maps are such an exhaustive accounting of ideas that they are difficult to translate into teaching practices.

The science education community, having realized that the field lacked documents that simultaneously filled the roles of setting standards for student learning as well as having instructional utility, has adopted a different perspective on the content contained in the NSES. Researchers are now developing tools called Learning Progressions, which are research-based sequences of conceptual understandings that map the terrain from naïve ideas and misunderstandings to scientifically accepted explanations (National Research Council, 2007). LPs are currently in development using multiple approaches to identify developmental sequences in many grade spans in a variety of content areas.

While LPs representing correct student understandings are important steps in the right direction, the science education community must still determine how we are going to make these tools into resources that improve teaching and learning. Exhaustive lists of learning goals do help teachers to know what students are to learn, but are difficult to implement in the absence of concrete information that will help teachers to identify these ideas when students share them during class, as well as concrete strategies for helping students develop more sophisticated scientific understandings (National Research Council, 2001).

Yet LPs have great potential to be invaluable teacher preparation and professional development tools since they contain information regarding knowledge of student ideas and how students learn, and – in some cases - suggestions for strategies or actions to help students learn. LPs thus could serve as tools to increase teacher knowledge by laying out how concepts are structured and interrelated, the sequences in which students learn, and suggestions for instruction. In order to serve that purpose, however, current formats for LPs need to include examples of student responses for each level, as well as instructional strategies and feedback suggestions to help students proceed up the LP.

Working continuously with such a learning progression could help teachers learn the structure and relationship of students' understandings about a particular concept, teaching strategies to elicit student thinking, example responses, and targeted feedback to students, showing them the steps they need to take to reach the steps on the learning progression, i.e. learning goals (D. R. Sadler, 1989). This process has been called a 'feedback loop' (National Research Council, 2001), and has been shown to have a significant effect on student learning (Black & Wiliam, 1998).

This paper describes one potential use of LPs as tools to support teacher development. It will set forth requirements for LPs to serve this purpose, will suggest designs for the development, validation, and study of LPs for teacher development, and will explore challenges of that research agenda in an example of ongoing research.

Learning Progressions

LPs are defined as "descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time" (NRC, 2007, p. 205). At least two approaches have emerged for the construction of learning progressions.

Type 1 Learning Progressions

The first type is a progression of *correct* ideas based on standards documents or consultations with scientists, articulated across grade spans (American Association for the Advancement of Science, 2001; Catley, Lehrer, & Reiser, 2005). LPs of this type identify how big ideas in science (American Association for the Advancement of Science, 1993) develop across different grade spans. This kind of learning progression tends to include an accounting of how scientifically accurate student ideas should develop during instruction; examples have been created in the domain of evolution (Catley et al., 2005) and modern genetics (R.G. Duncan, Rogat, & Yarden, in press). Type 1 LPs are most often used as guides for curriculum development.

Type 2 Learning Progressions

A second approach to LPs entails maps of student ideas bordered on one end by naïve ideas about the natural world that students hold prior to instruction and on the other end scientifically accepted explanations. The middle is occupied by intermediate understandings. This type of learning progression may be validated by two approaches.

The first approach employs the microgenetic method, a qualitative analysis which involves taking multiple, frequent samples of student knowledge at the time at which development is occurring for the purpose of identifying individual learning trajectories as well as average development of a concept (Siegler, 1997; Siegler & Crowley, 1991). For example, a researcher might interview students multiple times throughout a school year to explore how their ideas develop as a result of instruction.

The second approach relies upon quantitative analyses to validate the contents of the progression. Building upon the types of questions contained in conceptual inventories (Hestenes, Wells, & Swackhamer, 1992) and distractor-driven multiple choice items (P. M. Sadler, 1998), this approach explores the development of a latent construct by creating a preliminary learning progression or 'construct map' (Wilson, 2005), and then developing an assessment to validate it. The validation progress is conducted through the creation an assessment consisting of a set of ordered multiple-choice items (OMC's) in which each possible choice is linked to a level on the progression. The assessment is then given to a sample of students, and individual student performance is compared across items; if students select answers linked to a given level of the progression for most of the items, it can be inferred that student has reached that level of understanding (D. C. Briggs, Alonzo, Schwab, & Wilson, 2006). Furthermore, Item Response Theory (Lord, 1980) can be used to model the probability of correct responses for each item as student proficiency develops. A feature of this second type of learning progression is that it includes students' common prior ideas and errors; examples of these LPs already developed include those for student understanding of the position of the earth in the solar system (D. C. Briggs et al., 2006), sinking and floating (Nagashima, Brown, Fu, Timms, & Wilson, 2008), and force and motion (Alonzo & Steedle, in press).

Other Applications of Learning Progressions

While both Type 1 and Type 2 LPs commonly define and describe student learning, they are also being developed to track the development of teacher knowledge itself. Ongoing projects focus on scientific model building and argumentation (Ravit Golan Duncan, 2008), and science teachers' strategic knowledge (Talbot, Briggs, & Otero, 2009).

LPs have also been combined with classroom assessments for diagnostic purposes. One example among many produced by Researchers at the Berkeley Evaluation and Assessment Research Center (BEAR) is a map of student understanding of the perspectives of chemists, which is accompanied by assessment questions, a scoring rubric, and exemplars of student work (Claesgens, Scalise, Wilson, & Stacy, 2009). Such a suite of materials goes beyond an accounting of student ideas by providing additional tools to help students to identify levels of student understanding.

In addition, LPs have been developed to guide the development of student understanding through formative assessment. The Stanford Education Assessment Laboratory (SEAL) developed a series of formative embedded assessments linked to a learning progression for students' relative density-based explanations for sinking and floating. However, an experimental study of the effects of the progression and accompanying formative assessments yielded mixed results (Yin et al., 2008). An implementation study of how teachers used the formative assessments indicated that teachers who were able to combine the information contained in the learning progression with a variety of instructional strategies and as a basis for providing informational feedback had students with higher learning gains (Furtak et al., 2008).

The result of this study raises questions about the utility of LPs and assessments in isolation from concrete feedback strategies to scaffold teachers' ability to 'move students up' the learning progression. In fact, during the SEAL study, a teacher contacted the research team asking for feedback suggestions for students with the idea that air alone causes things to float; it took several days of thinking and research for SEAL to produce a list of ideas for this teacher, calling into question the feasibility of asking teachers to come up with responses on the fly (Ruiz-Primo, Furtak, Ayala, Yin, & Shavelson, in press).

Knowing the means by which students learn concepts is important, but without the tools to help teachers recognize and act upon those ideas, the means by which the information LPs contain will be translated to practice is unclear. Furthermore, while developing LPs to track teacher development is important to better understanding how constructs such as pedagogical content knowledge (PCK) develop, an important intermediate step is to develop LPs of student understanding that will scaffold the development of the knowledge and practices that teachers need to for effective science instruction.

Given the aforementioned challenges presented by previous studies employing LPs in the classroom, it follows that LPs could be complimented with additional information to make them not only maps of how students' knowledge develops through instruction, but also maps of the instructional strategies that will help teachers develop in their abilities to effectively teach a given concept.

Requirements for LPs as Tools for Teacher Development

Shulman (1986) argued that teachers' PCK, or content knowledge *for* teaching, includes understandings of students' common prior ideas and "knowledge of the strategies most likely to be fruitful in reorganizing the understanding of learners" (p. 9-10). Therefore, if LPs are accompanied by teaching strategies tailored to students' ideas, they can serve as maps of the knowledge necessary to teach a particular scientific concept, help teachers learn the structure and relationship of students' understandings about a particular concept, and provide targeted suggestions for instruction.

This section suggests four design criteria for LPs in the service of teacher development. These LPs should include an exhaustive accounting of student ideas, assessment prompts that can help teachers find out what students know, provide age-appropriate examples of student ideas, and present a set of feedback strategies for each level in the progression. The sections below draw presents each criterion in more detail, supported by the research base in science teaching and assessment.

Criterion 1: Accounting of student ideas as they develop

LPs for Teacher Development should combine features of Type 1 and Type 2 LPs to include not only the multiple and sequential *correct* understandings that build to the big ideas in science across multiple grade spans, but also the development of student ideas on a single construct beginning with alternative conceptions and ending with scientifically accurate understandings. However, the way that a 'single construct' is defined is an important point of clarification. The AAAS's *Atlas of Science Literacy,* for example, attempts to map all of the interrelated knowledge necessary to understand certain big ideas in science, such as cellular functions (American Association for the Advancement of Science, 2001). However, the maps integrate so much interrelated knowledge that their instructional utility is limited, and student progress on a single construct cannot be accurately determined (see Wilson, 2005).

Initial development of such LPs should begin with extensive mapping of student thinking in a conceptual domain on the basis of prior research and new empirical studies, and involve continuous empirical validation of a preliminary progression through fine-grained analyses of student work and responses to OMC items on assessments (D. Briggs, 2008; Wilson, 2005).

Criterion 2: Formative assessment prompts to elicit student thinking.

Before teachers can take instructional action on the basis of student ideas, they need high-quality assessment prompts to elicit student thinking (Pellegrino, Chudowsky, & Glaser, 2001). Such prompts should entail a variety of formats and be linked to the learning progression to explicitly draw out the ideas that appear in the progression, so that teachers can identify the levels at which students are.

To serve the purpose of advancing student learning, assessment prompts linked to the learning progression should be formative in nature; that is, intended to provide a basis for feedback rather than being graded. The prompts should be generative in nature, inviting students to share a range of ideas.

Formative assessment prompts can be described along a continuum between pre-planned, curriculum-embedded prompts and spontaneous, ‘on-the-fly’ prompts, shown in Figure 1 (Furtak & Ruiz-Primo, 2008; Ruiz-Primo & Furtak, 2007; Shavelson et al., 2008).

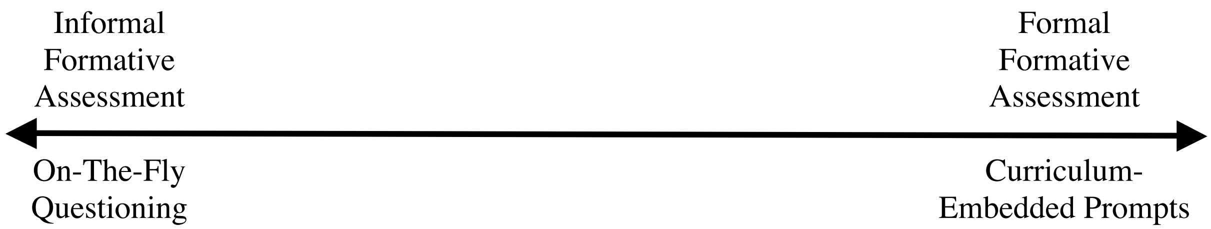

Figure 1. *Continuum of formative assessment* (Shavelson et al., 2008)

In this sense, *informal* or “on-the-fly” formative assessment takes any interaction in class (i.e., one-to-one interactions, small group, or class discussion) as an opportunity to gather and elicit information about students’ level of understanding in order to provide appropriate and timely feedback. The teacher can respond immediately to students’ ideas, making possible the completion of multiple feedback cycles in a relatively short period of time (Ruiz-Primo & Furtak, 2006). At the other extreme lies *formal* formative assessment, consisting of pre-developed assessment prompts that help to guide the collection of students’ conceptions. A formal formative assessment prompt is a question or task that is aligned with learning goals and serves the purpose of eliciting students’ conceptions as a basis for teachers to make instructional decisions to reduce the gap between learning goals and students’ present state of understanding.

Formative assessment prompts need to be well designed in order to make student thinking explicit. An assessment prompt specifies what the student will say, do, or create to provide the necessary evidence what students know. The design of the assessment prompts should start with clarity about the knowledge or skills that are linked to the instructional goals at hand. This specification will lead to the design of the means by which students’ knowledge or skills will be elicited, starting with the task, moving to the characteristics of the students’ responses that should be captured, and ending with the strategy used to judge the quality of the responses.

Criterion 3: Age-appropriate examples of actual student responses

Current versions of LPs provide summaries or categories of student ideas; however, these are often phrased in general terms and do not reflect the ways that students speak about scientific ideas. For example, where density is concerned, students often will speak of objects being heavy for their size, or about an amount of mass being ‘spread out’ in a particular amount of space. A learning progression might summarize these ideas as naïve understandings of density, but without examples of these kinds of student statements, teachers may have difficulty in properly categorizing them.

To facilitate teachers identifying student understanding on a range of formative assessment prompts (and informal observations and discussions), age-appropriate responses for each level of the learning progression should be provided. These student ideas could be collected in the process of validating the learning progression itself with a sample target student audience.

Criterion 4: Suggested feedback strategies to move students between levels
The key element of LPs for teacher development is the inclusion of feedback strategies to help teachers move students between levels. Lower levels of feedback, such as providing evaluation ('correct/incorrect') or acknowledging students' effort ('I can tell you're working very hard') are not as effective as feedback that contains information about student performance and how to meet learning goals (Hattie & Timperley, 2007).

Informational feedback strategies can take many forms, from questions asked by the teacher, to integrating additional descriptions or representations of particular ideas, to revised lessons plans, to re-teaching entire units when students do not understand. Unfortunately, these probing questions, strategies, activities, and representations are difficult for most teachers to come up with on-the-fly; in fact, Talbot et al. (2009) found that only the most sophisticated teachers were able to respond with informational feedback in unexpected situations.

LPs in effect anticipate for the teacher the different kinds of responses that students will give during instruction, and thus are an ideal starting point for suggesting informational feedback that will help advance student learning. If LPs are developed in cooperation with practicing teachers, these teachers can help suggest different strategies to include on the learning progression. Furthermore, classroom-based studies can explore the differential effectiveness of various strategies on helping students advance in their thinking, thereby providing another source of evidence for the validity of LPs for teacher development.

Research Designs

Given the close link between LPs for teacher development and instructional practices, it is essential that researchers and practicing teachers develop them in close collaboration. Researchers can bring the necessary expertise to develop a preliminary version of the learning progression, and then can work with teachers to develop OMC's to validate the progression, collect student work to explore individual student learning trajectories, develop and pilot assessment prompts, collect examples of student work, and develop, enact and refine feedback strategies.

LPs for teacher development will be validated not only through careful analyses of student thinking, but also through classroom-based studies that explore the relative effectiveness of different teaching approaches in helping students adapt their prior ideas into more scientific understandings. This means that research into LPs for teacher development must be conducted in two phases; first, the development and validation of the learning progression, and second, the empirical exploration of its impact on student development and student learning.

The first phase, LP development and validation, necessitates an iterative process. It should begin with a hypothesized progression of student development, which is then used to develop a series of assessment prompts that will bring evidence to bear on its validity. The hypothesized LP can be developed on the basis of prior research into student thinking. At the same time, the hypothesized LP should be circulated among teachers to determine its instructional utility, and should be employed in classroom practice for the purpose of collecting sample student responses and teacher feedback ideas. After each cycle of data collection, the hypothesized LP should be revised and retested, and the final product should contain all four criteria.

Once the LP has been validated, the second phase of research will entail empirical exploration of its impact on professional development and student learning. This research should take place with a new set of teachers who were not involved in the development of the LP, and should employ a multiple-method approach (Smith, 2006). Data collection related to both student learning and teacher development should be collected; student learning should be tracked using pre-post and formative assessments linked to the LP, and teacher development should be triangulated through interviews, classroom observations, and questionnaires administered over a course of multiple teachings of the relevant content.

Challenges in Developing LPs for Teacher Development

The research area of LPs is still emerging, and the field yet to agree on the purpose or methodological approaches for developing Type 1 and 2 LPs. Furthermore, the application or 'use' of both types of LPs for multiple audiences – students, teachers, administrators, curriculum developers, assessment designers, and so on - is still in its infancy. For the purpose of illuminating an example of current research to inform other projects, this section will briefly introduce an ongoing research project, will present the LP currently being developed and validated, and will then explore challenges faced in the project, and how they have been addressed.

Developing a LP for Teacher Development: The Daphne Project

Charles Darwin's theory of evolution through the process of natural selection is the unifying framework for Biology (Dobzhansky, 1973); nevertheless, students often have a difficult time understanding it (Bishop & Anderson, 1990; Ferrari & Chi, 1998). Studies indicate evolution instruction in high school has been "absent, cursory, or fraught with misinformation" (Rutledge & Mitchell, 2002, p. 21), and students commonly have misconceptions about how populations change over time (Anderson, Fisher, & Norman, 2002; Rudolph & Stewart, 1998; Shtulman, 2006).

The goal of the Daphne Project was to create a LP for Teacher Development by working with a department of biology teachers to explore and validate a hypothesized learning progression. It was hypothesized at the outset of the study that engaging in the process of developing this LP would help teachers be better prepared to recognize and act upon student thinking, which in turn would impact student learning by closing what has been called a 'feedback loop' in the formative assessment process (Black & Wiliam, 1998; National Research Council, 2001).

The Daphne Project, supported by a Research Fellowship from the Knowles Science Teaching Foundation, is currently in the second of three funding years. The first year (Y1) involved development of a Type 2 LP based on prior research into students' ideas about natural selection, as well as piloting of the pre-posttest and formative assessment activities designed to collect student responses. The second year (Y2) involved further development and use of a revised LP with a group of biology teachers. The same group of teachers will continue to refine and use the LP in the third year (Y3), the outcome of which will be a 'deliverable' LP for teacher development whose impact on teacher knowledge and student learning will be explored in a future 'Phase 2' research project. A project timeline for these stages is show in Table 1.

Table 1. Daphne Project Timeline

Project Year	Phase	Project Activities
Y1 2007-2008	1	Develop hypothesized LP and instruments; collect student responses to instruments; analyze data and make revisions to all instruments
Y2 2008-2009		Pilot 1 of LP and instruments in regular classroom settings; analysis of results and revision of LP and instruments
Y3: 2009-2010		Pilot 2 of revised LP, instruments, and assessments; collect classroom strategies and develop 'deliverable' LP
Y1 etc.	2	Study impact of deliverable LP on teacher knowledge and student learning at new school sites

Participants

Seven biology teachers and their students (N=341) at Springfield High School participated. Springfield, considered among the top high schools in the state, is an ethnically and socioeconomically diverse school near a large city in the western US. The seven teachers represented a wide range of backgrounds and experience, ranging from a second-career student teacher to a 15-year veteran Biology teacher. Students in the participating teachers' classes were enrolled in three levels of biology: 9th grade pre-International Baccalaureate Biology, 10th grade General Biology, and 10th grade English Language Learner (ELL) Biology.

Professional Learning Community

The six teachers and the lead researcher participated in meetings every two to three weeks during the 2008-2009 for the purpose of establishing a professional learning community (PLC) centered around the theory of natural selection, approaches to teaching it, the LP, and its accompanying formative assessments and feedback strategies. Sessions involved discussing approaches to teaching evolution, predicting and exploring students' ideas about natural selection, categorizing student work, and discussing different instructional approaches given student ideas represented in the LPs.

LP for the Teaching of Natural Selection

In the past two decades, a number of student ideas about natural selection have been identified (Bishop & Anderson, 1990; Dagher & Boujaoude, 2005; Ferrari & Chi, 1998; Geraedts & Boersma, 2006; Shtulman, 2006). By organizing students' ideas in a linear fashion, progressing from simple to complex, moving away from the ideas of Lamarck toward those of Darwin (Geraedts & Boersma, 2006; Shtulman, 2006).

The Daphne Project began with a draft LP, which was developed by the research team based on existing literature and a Y1 pilot. This LP, called a 'map of student ideas' in the study, was brought to the PLC as an organizer for instruction during Y2 of the study. Its four basic levels start with unclear or undifferentiated ideas (1), then the classic 'misconceptions' about natural selection (2), a blending of these 'misconceptions' with terms used to describe natural selection (3), and the classic 'correct' understanding of natural selection (4). Although a number of 'flavors' of student misconceptions about the origin of new traits exist, it is unclear in the research base if any kind of progression exists for the multiple common misconceptions students have about natural selection (e.g. 'environmenally-induced' versus 'anthropomorphic' changes). The draft or hypothesized LP contained the list of student ideas, example student responses collected during the Y1 pilot study, and feedback brainstormed by the Daphne Project Team (Table 2).

Table 2. Map of Student Ideas about Natural Selection

Category	Description	Example Student Responses	Suggested Feedback
Natural Selection	Changes in populations and species occur gradually over time, due to the fact that all organisms produce more offspring than are able to survive. These offspring vary among themselves, and some of the variation can be passed on to the next generation. Thus, those offspring that are able to survive will happen to be those best suited their environments. Over many generations, accumulated changes will lead to changes in isolated populations and, in some cases, the generation of new species.	Natural selection occurs when the environment of the organism changes for whatever reason. The first graph shows that the length of the beaks of the birds were relatively spread out from 8-11.5 mm. However, the second graph shows a sharp increase in beaks that were about 10.1-10.7 mm long. Given the information about the drought, these graphs show that the birds with slightly larger beaks had a little bit more of an advantage over the other birds during the drought. This shows natural selection because it shows that those that had the ability to cope with the changing environment were more likely to successfully reproduce.	
Blended (Natural Selection and Need-based change)	Students base explanations on features of natural selection as well as need-based change; often include descriptions of the principles of 'survival of the fittest' and/or 'non-inheritance of acquired traits' in an otherwise non-selection description	The moth's predator would have killed the *Typica* and in order to survive the genes changed making it harder for the birds to see them. The moth needs to survive and to do that it blends in with the bark. If the bark gets darker the moth needs to go through some genetic mutations to match the bark. You can't have a white moth on dark bark, it wouldn't survive.	Tell students that scientists have found that environments do not cause individuals to change, so then how does the variation come about?
Need-based change	<u>*Anthropomorphic*</u>*:* Changes arise from deliberate acts by the parent organism or the species as a whole <u>*Environment*</u>*:* Changes emerge because of environmental conditions; i.e., the environment causes changes in organisms to occur	The moths become darker because of bark The moths would change their color to a darker color to blend in with the dark bark. When food and other resources become scarce, the finches develop over a period of time different beaks to eat the hard seeds that seem to be abundant. Animals mutate to fit in with their natural surroundings. So becoming darker helps to keep them in camouflage.	Ask students, 'How do changes in the environment cause changes in individuals?' Ask students, 'How do mutations in an individual get passed on to their offspring?'
Unclear/ Other	Ideas that are confusing, ambiguous, or do not fall into other categories	When the world is evolving and changing populations can't stay the same because they won't be able to survive. So this is why species evolve as an environment changes species need to adapt to it. Populations change over time because there has to be a balance between species. If there is no balance some species will over populate, which can cause problems to the whole world. If populations stay at a balance then everything is equal and everything will	Get at what students think is the underlying mechanism for change by asking what students mean by 'adapt,' etc. Ask students where changes in populations come from

	function right. The cause of natural selection was an increase in population. Natural selection helps a species to adapt to obstacles and survive.

This draft learning progression – which the project team called a ‘map of student ideas’ – included three of the four criteria for LPs for teacher development, the fourth criterion of assessment prompts to elicit student thinking were provided to teachers separately. Some of these assessment prompts are shown in Figure 2.

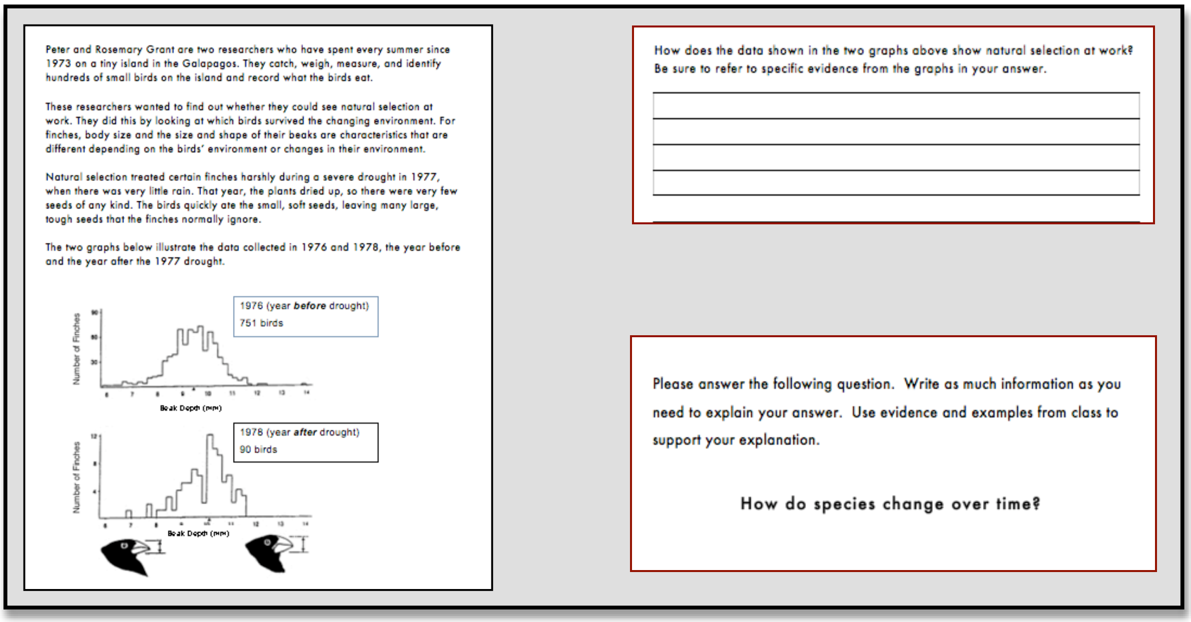

Figure 2. Sample draft assessment prompts for the draft learning progression.

Sources of Data

Multiple data were collected during Y2 of the study to bring evidence to bear on the validity of the draft LP, as well as to begin to explore the ways the teachers used the LPs during their natural selection unit. Students completed a version of the CINS (modified to reduce the language load, but with no content changed) pre-post unit, as well as a series of formative assessments linked to the LP. Teachers participated in a pre-, mid-, and post-interviews regarding their use of the LP and their knowledge of student ideas, and also answered questions on and predicted student responses to the CINS pre and post. In addition, researchers facilitated a representation of the teachers’ knowledge about how to teach natural selection by the group of teachers (Loughran, Mulhall, & Berry, 2004). At least one instance of each teacher’s implementation of the formative assessments was videotaped for analysis of how the prompts worked, the ideas students shared, and the types of feedback the teachers provided in response to those ideas; these videos were watched alongside researchers in stimulated recall interviews at the mid-interview. These measures of teacher knowledge were given not only to explore teachers’ use of the LP in practice, but also to begin to understand which measures might help triangulate changes in teachers’ knowledge in response to the LPs in Phase 2 of the study. Since

PCK has proven to be a difficult construct to measure with a single instrument (Hill, Ball, Blunk, Goffney, & Rowan, 2007), changes in teachers' knowledge will be triangulated with these multiple sources of data in future iterations of the study. The research questions for the study and associated sources of data are listed in Table 3.

Table 3. Daphne Project Research Questions and Sources of Data.

Research Question	*Sources of Data*
How does high school student thinking about Natural Selection develop over the course of a unit? How might that development be represented in an instructionally useful manner?	Pre-Post CINS Formative Assessments Classroom Videos
How does a representation of these student ideas impact teacher development?	Predicted responses to Pre-Post CINS Natural Selection Unit representation Classroom Videos PLC Videos Pre, Mid, and Post-Interviews

Although data from the Y2 are still in the process of being analyzed, our experiences during Y1 and Y2 of the study have presented us with a number of challenges in this new research. For the remainder of the paper, we will present the challenges we faced and the ways in which we addressed – or plan to address – them.

Challenges Faced in Creating LPs for Teacher Development

In the course of the research described above, we - the Daphne Project research team - have encountered a number of challenges related to the development and validation of the learning progression for the purpose of building teacher knowledge. The challenges are summarized in Table 4; each of these challenges is described in detail below.

Table 4. Summary of challenges faced in developing LPs for Teacher Development

Category	Challenge Faced
LP Development	Teasing out multiple constructs in LP Dealing with changing LP during study Combining Type 1 and 2 LPs
LP Enactment	Viewing student competence as a continuum Anticipating feedback Scaling up

Challenge 1. Teasing out multiple constructs in the LP

At the outset of the study, we assembled the draft Type 2 LP – which we at that point called a 'map of student ideas' – based on prior research, shown in Table 2. That draft LP showed what we believed to be a single construct, moving from student preliminary ideas that are the common student alternative conceptions about natural selection toward a complete explanation of Natural

Selection. Based on this draft LP, we decided to work with the Conceptual Inventory of Natural Selection (CINS) as our pre-post measure, as it was developed using the same literature base that helped to generate the draft LP, and the questions mapped more or less cleanly onto the levels represented in the draft LP.

However, during Y2 of the study, we realized that that what we had represented as a singular construct actually contained multiple confounded constructs. The top level of the LP, which we had titled ‘natural selection’, actually contained information not only about the origin of new traits within a population – a classic area for alternative conceptions - but it also contained information about population dynamics and ecology. According to Wilson (2005), construct maps or LPs should contain only a single construct to serve as a valid and reliable measurement tool.

Teasing out these ideas was made difficult by the multiple and interacting constructs present in natural selection; for example, student understanding of natural selection draws on foundational understandings of ecology, population dynamics, and genetic inheritance (Catley et al., 2005). During Y2 of the study we spoke with more biologists and became aware that a barrier to student understanding of the genetic basis of evolutionary change is students’ difficulty in recognizing the role of random molecular processes in biology (Garvin-Doxas & Klymkowsky, 2008; Kelemen & Rosset, 2009). Furthermore, students’ impressions of genetic mutations are usually that they are always harmful, given the use of the word ‘mutation’ in popular media, comic books, and films (7-Eleven’s ‘Mutant Berry’ flavored Slurpee as a promotional item for the ‘X-Men’ is just one such example).

Given this additional information, we realized that once students understand that changes in individual organisms are results of random genetic changes, they often view mutations as being creative or deleterious only before they understand that there are many kinds of mutations, most of which have no effect, and only a few of which compound over time to lead to the generation of new genetic traits (Klymkowsky, Furtak, Cooper, Garvin-Doxas, & Gonzales, draft). If those traits are beneficial in some way, individuals with those traits will be more likely to survive and have offspring, thereby changing the composition of the population over time. The difference between these two elements of natural selection suggested addition of an additional category of student thinking to our hypothesized LP to reflect how students’ incomplete understanding of the role of genetics in the generation of new traits could be integrated into their alternative conceptions.

As coders, we have found these constructs to be independently represented in student responses. For example, students commonly combine elements of a correct description of selective force with a need-based source of new traits, as in the sample student response below:

> Species change over time because of where they live. For example, a school of fish gets thrown on land because of a tsunami, the fish would more than likely die but one femail [sic] fish lays her eggs on land. When those fish are born, they will have no water around so they have to adapt to land.

This response combines ideas related to the origin of traits within a population – in this case, an organism needing to adapt to a changing environment – with ideas related to reproduction and selection – certain fish dying and others successfully reproducing. Scoring a response like this according to the draft LP was almost impossible, since ideas relating to natural selection and alternative conceptions are both present.

We addressed this challenge during our analysis of these and other items during our Y2 revision phase. Rather than trying to fit these diverse ideas into one LP, we broke the LP apart into two separate constructs, one for the origin of new traits, and the other for the selection of traits within a population. While we acknowledge that other constructs are also present to assemble a complete conceptual understanding of natural selection, we are currently working with only these two constructs, illustrated in Figure 3 below.

<table>
<tr><th colspan="4">Origin & Inheritance of Traits</th></tr>
<tr><td rowspan="3">Genetic basis</td><td colspan="3">Random mutation or recombination of genes</td></tr>
<tr><td colspan="3">Undefined mutation</td></tr>
<tr><td colspan="3">Need-based mutation</td></tr>
<tr><td>Need-based change</td><td>Anthro-pomorphic</td><td>Environmentally Induced</td><td>Eugenic</td></tr>
<tr><td colspan="4">No Explicit Mechanism</td></tr>
<tr><td colspan="4">Construct Not Present</td></tr>
</table>

<table>
<tr><th colspan="2">Selective Force</th></tr>
<tr><td colspan="2">Selective Force</td></tr>
<tr><td>Reproduction</td><td>Selection</td></tr>
<tr><td colspan="2">Undefined Selective Force</td></tr>
<tr><td colspan="2">Construct Not Present</td></tr>
</table>

Figure 3. Revised Type 2 LPs for Natural Selection

Using these revised constructs, the example student response shown above would be scored as 'environmentally induced' on the origin and inheritance of traits construct, as well as a description of selective force including ideas relating to reproduction and selection by the environment.

While the Daphne Project team feels confident these revised Type 2 LPs more accurately represent the student thinking we observed in the study, we also realize that future iterations of the project may yield additional Type 2 LPs broken out of the original draft LP; for example, Anderson et al. (year) also discuss students' common misunderstandings of the term 'fitness,' which we integrated into the 'Selective Force' LP. It is entirely within the realm of possibility that this construct may also be confounded and may need to be broken into smaller constructs.

Challenge 2: LP as a moving target during Phase 1 development

Unfortunately, we noted the shortcomings in our draft LP early in Y2 of the Daphne Project, leading us to Challenge 2, which was dealing with an ongoing research project while the very tool we were creating was necessitating change right under us. The changing LP forced us to quickly discuss whether or not we were willing to change the LP wholesale mid-year, or whether or not we should stay with the draft version. The shortcomings were not limited only to the 'map of ideas' on the LP, either; we also realized that some of the formative assessment prompts themselves also confounded the constructs we had now divided. For example, take the multiple-choice item we intended for a formative assessment and discussion prompt shown in Figure 4.

Typica

Carbonaria

(Text and images retrieved from http://www.millerandlevine.com/km/evol/Moths/moths.html)

The difference in the number of light and dark-colored peppered moths in the population can be explained by:

A. Birds, which were more likely to eat the light-colored moths than the dark-colored moths. The darker moths were then more likely than the lighter moths to survive and reproduce.
B. Industrial pollution, which caused the bark of trees to darken over time. This caused the coloring of moths to become darker over time
C. A change in the color of bark, which caused the genes in a moth to mutate to produce a darker color. Then the moths with the mutation of a darker color were more likely to survive.
D. The moths themselves, which were able to make themselves become a darker color to match the bark of the trees, which had been darkened over time by industrial pollution.

Figure 4. Moth Multiple-Choice Formative Assessment

Upon our analysis of student responses, we realized that the question above confounds constructs regarding the origin of variation/new traits with differential survival and reproduction. For example, responses B-D, the distractors tied to classic student alternative conceptions, each cite a mechanism for the origin of variation (e.g., the color of bark) combined with information about differential survival (e.g., the moths with the mutation of a darker color were more likely to survive). In contrast, response A, the correct response, includes only a selective mechanism, and does not address how the dark color of the moths originated.

Similarly, when we turned to interpret the CINS, our prepost measure during Y2, we again founds multiple constructs confounded. Take as an example item 11, which asks students questions about Venezuelan Guppies (Figure 5).

11. Assuming a pair of guppies had plenty of food and space and no predators, what would happen if were they were placed in a large pond?
    a. The group of guppies would grow slowly, as guppies would have only enough babies to keep the group the same size.
    b. The group of guppies would grow slowly at first, then would grow quickly, and thousands of guppies would fill the pond.
    c. The group of guppies would never become very large, because only living things such as insects and bacteria reproduce that way.
    d. The group of guppies would continue to grow slowly over time.

Figure 5. CINS Question 11 (Fisher, Anderson, & Norman, 2002)

Again, this item confounds the origin of variation (responses A, C, and D) with selective force (response B), which makes difficult the interpretation of student responses on a single construct.

In the end, we did not change the map of student ideas while Y2 of the study was in progress, since at that point we were not analyzing data systematically, and the LP still contained an important distinction between ideas that we summarized as ‘need-based,’ including anthropomorphic and environmentally-induced ideas of change, as well as blended ideas (need-based ideas combined with correct ideas).

We addressed the challenge of the now-outdated LP and our flawed formative assessments by continuing our dialogue with the teachers, being explicit in relating that we were in the developmental phases of a new research agenda, and that any an all items we were working with were still only hypotheses that could be modified as we collected additional evidence. We invited the teachers to help us note areas in which the LP seemed inaccurate or not fully descriptive of the ideas we were noticing as the study continued.

In one instance, two teachers met together to explore student work and identified what they began calling the ‘Eugenic’ misconception about natural selection, as described by Teacher 2 below:

> We found as we read them that there's all this eugenic idea that the animals are choosing like we want our kids to be better. So not that they're actually changing but they're choosing who their mate is because they want their kid to be like this so they choose their mate. So then we added on E [to the moth multiple-choice formative assessment], that they choose their mate because they want the next generation to look a certain way.

In this way, teachers in the study were actively involved in identifying new nuances in student thinking and suggesting additional hypothesized levels to the LP.

Furthermore, activities created as a group during Y2 of the study helped us to tease out more explicitly the confounded constructs. As we discussed the map of student ideas, teachers suggested that possible sources of student alternative conceptions might come from overgeneralizations based upon everyday observations of nature. For instance, students often believe that individuals can actively modify themselves, an anthropomorphic idea that may come from observations of humans changing themselves, or of chameleons that can change their color to match their surroundings (Kelemen & Rosset, 2009). Similarly, students also believe that the environment can ‘cause’ changes in organisms, such as those that change the color of their fur with the seasons (such as the arctic hare). To explicitly draw out these ideas, the teachers and researchers collaborated to develop a new formative assessment that would present these potentially confusing instances of changing organisms and ask students to explain why this was or was not an instance of a change in a population (i.e., natural selection), or a change in an individual organism, and to explain why they thought so (Figure 6).

5	Arctic hares live in the harsh environment of the North American tundra. In the winter, the hares have a white coat of fur, helping them to blend in with ice and snow. In the summer, their coat turns a gray-blue that matches rocks and plants in their environment.	☐ Natural Selection ☐ Individual Change *Why do you think so?*
6	Antibiotics are used to kill bacteria. Since the 1950's, however, scientists have noticed that some bacteria are not be killed by some kinds of antibiotics. When you take an antibiotic, all the non-resistant bacteria are quickly killed and the remaining resistant bacteria are then left to reproduce and thrive.	☐ Natural Selection ☐ Individual Change *Why do you think so?*
7	While some Canada Geese still fly south for the winter, other 'resident' populations of Geese live in the same place all year. Migration – the seasonal journey in response to food availability, weather, or habitat - is a learned behavior, so once the adults in a population stop migrating, their offspring will not also not migrate south.	☐ Natural Selection ☐ Individual Change *Why do you think so?*

Figure 6. Individual vs. Population Change Formative Assessment

Despite our openness with the teachers regarding our need for input and feedback at this developmental phase of research, we still received some pushback from the teachers. On several occasions, one or more teachers approached us to ask us why we were not telling them directly what they should do during the research project. From their perspective, 'research' usually was done from an experimental perspective, in which the teachers would all implement the same activities in the same sequence, and we would then determine the effectiveness of the intervention. We addressed these concerns by telling teachers that, since we were still developing the LP and associated assessment prompts, we did not yet feel comfortable to ask them to all do the same thing, and reminded them that the diverse data we collected during Y2 could help us determine the best course of action for all teachers to take during Y3 of the study.

Challenge 3. Combining Type 1 and 2 LPs

When we began the research, we brought only a draft Type 2 LP to the PLC, as described above. However, beginning with the pre-project interviews, some of the teachers described their organization of the natural selection unit based on Mayr's five facts and three inferences, a common way of describing the process of natural selection (Mayr, 1997):

> *Fact 1:* All populations have the potential to grow at an exponential rate.
> *Fact 2:* Most populations reach a certain size, then remain fairly stable over time.
> *Fact 3:* Natural resources are limited.
>> *Inference 1:* Not all offspring survive to reproductive age in part because of competition for natural resources.
>
> *Fact 4:* Individuals in a population are not identical, but vary in many characteristics.
> *Fact 5:* Many of the characteristics are inherited.
>> *Inference 2:* Survival is not random. Those individuals with characteristics that provide them with some advantage over others in that particular environmental situation will survive to reproduce, whereas others will die.
>>
>> *Inference 3:* Populations change over time as the frequency of advantageous

alleles increases. These could accumulate over time to result in speciation.

Mayr's account states in simple terms the stepwise understandings one needs to have, and the inferences one needs to make on the basis of those understandings, in order to grasp the whole of natural selection. The five facts and three inferences are also the basis of the description of natural selection in the *National Science Education Standards* (NRC, 1996). The Daphne Project team quickly realized that, given the way Mayr's account was being used by teachers and its contents, was essentially functioning as a simplified Type 1 LP (Table 5).

Table 5. Draft type 1 LP for Natural Selection

<table>
<tr><th>Level</th><th colspan="6">Description</th></tr>
<tr><td>5</td><td colspan="6">Populations change over time as the frequency of advantageous alleles increases. These could accumulate over time to result in speciation.</td></tr>
<tr><td>4</td><td colspan="6">Those individuals with characteristics that provide them with some advantage over others in that particular environmental situation will survive to reproduce, whereas others will die.</td></tr>
<tr><td>3</td><td colspan="3">Individuals within a population vary in many characteristics</td><td colspan="3">Many, but not all of these characteristics are inherited</td></tr>
<tr><td>2</td><td colspan="6">Not all offspring survive to reproductive age in part because of competition for natural resources.</td></tr>
<tr><td>1</td><td colspan="2">All populations have the potential to reproduce exponentially.</td><td colspan="2">Most populations reach a certain size, then remain fairly stable over time.</td><td colspan="2">Natural resources are limited</td></tr>
</table>

In this representation, Mayr's 'facts' about natural selection are listed on the same level to suggest they could be taught together or in close succession (Levels 1 and 3), and the steps above those facts that require 'inferences' from those facts (Levels 2, 4, and 5). The introduction of Mayr's representation into our study added another layer of complexity we had not anticipated. It introduced tension between these two representations of student thinking about natural selection.

Within our research group, as well as within the PLC, we began to grapple with the question of how all of these LPs we were creating were related to each other. As the mid-interview, we asked teachers directly how the two kinds of LPs were influencing what they were doing in the classroom. Two of the teachers have dismantled their units as they were currently taught and reconstructed their units based on Mayr's facts and inferences as represented in the Type 1 LP, building concepts up to the understanding of natural selection rather than starting with the theory and going through different pieces of evidence for the theory, as they had before. Both of these teachers commented that the Type 1 LP was useful because it helped them to see how concepts from genetics and population ecology, traditionally the units that come before and after the natural selection unit, also needed to be taught as part of evolution.

Prior to the teaching of the natural selection unit, the teachers and Daphne Project team matched student work with the levels on the Type 2 LP, and created a series of activities designed to elicit and address the student ideas it includes. Observations of and interviews with the teachers during the unit enactment suggest that these tools have motivated several changes in the teachers' practices, as well as their descriptions of their approaches to teaching natural selection. For example, as a preassessment, two teachers asked students to summarize in writing everything they knew about natural selection, and worked together to categorize the students' responses according to the levels. The groups for the unit were then constructed to contain students with diverse ideas. A third teacher purposefully chose activities that precisely address the ideas contained in the Type 2 LP.

At the end of the unit, we asked teach teacher directly how the two kinds of LPs interacted during their teaching. Most of the teachers felt that both of the representations were important, using the Type 1 LP as a guide for the scope and sequence of their Natural Selection units, and then using the Type 2 LP to help them identify and address student thinking during the unit.

> I think that they were both equally useful, because one allowed me to organize the unit, and the other one allowed me to organize specific lessons, and get at specific misconceptions that I hadn't really thought of before…

Teachers' individual comments mirrored teachers' reactions in PLC meetings, during which teachers argued that both the Type 1 and 2 versions of the LP are necessary to organize and conduct instruction. The first kind of LP is necessary for setting learning goals and determining the scope and sequence of the unit, and the map of incorrect ideas (Type 2) helps them to elicit and take action upon misconceptions as they arise.

Integrating both types of LPs reflects a solution to the 'either-or' dilemma, a false dichotomy that advances one type of LP over the other, as described by Shepard (2009). In response to this feedback from teachers and following conversations with other researchers, we decided to combine the Type 1 and Type 2 LPs into a single representation, with the Type 1 LP of Mayr's five facts and three inferences spanning the top of the figure, and with the relevant Type 2 LPs – to date we have developed only two – supporting the development of concepts represented in the Type 1 LP (Figure 7).

In future iterations of the project, this combined Type 1 and Type 2 LP will be used as an overall organizing framework, whereas the Type 2 LPs will be further explicated according to the design criteria presented earlier in this paper, as shown in Appendix A.

LP for Natural Selection

Unit Progress →

Facts 1, 2 & 3	Inference 1	Facts 4 & 5	Inference 2	Inference 3
All populations have the potential to reproduce exponentially.	Not all offspring survive to reproductive age in part because of competition for natural resources.	Individuals within a population vary in many characteristics	Those individuals with characteristics that provide them with some advantage over others in that particular environmental situation will survive to reproduce, whereas others will die.	Populations change over time as the frequency of advantageous alleles increases. These could accumulate over time to result in speciation.
Most populations reach a certain size, then remain fairly stable over time.		Many, but not all of these characteristics are inherited		
Natural resources are limited				

Source: Mayr, 1997

<table>
<tr><th colspan="4">Origin of New Traits</th></tr>
<tr><td rowspan="3">M</td><td colspan="3">Random mutation</td></tr>
<tr><td colspan="3">Undefined mutation</td></tr>
<tr><td colspan="3">Need-based mutation</td></tr>
<tr><td>NB</td><td>Anthropomorphic</td><td>Environmentally Induced</td><td>Eugenic</td></tr>
<tr><td colspan="4">No Explicit Mechanism</td></tr>
<tr><td colspan="4">Construct Not Present</td></tr>
</table>

<table>
<tr><th colspan="2">Selective Force</th></tr>
<tr><td colspan="2">Selective Force</td></tr>
<tr><td>Reproduction</td><td>Selection</td></tr>
<tr><td colspan="2">Undefined Selective Force</td></tr>
<tr><td colspan="2">Construct Not Present</td></tr>
</table>

Figure 7. Combined Type 1 and Type 2 LPs for Natural Selection

Challenge 4. Viewing student thinking as a continuum
A hypothesized advantage of using LPs for teacher development is that they will help teachers move from a traditional view of student thinking as correct or incorrect toward viewing student ideas as a continuum based on students' prior experience.

Prior to the study, the majority of the teachers organized their natural selection units by presenting Darwin's theory, and then completing a series of activities that illustrated the evidence for the theory. Three of the six teachers identified one or more common student misunderstandings of natural selection or evolutionary theory; this did not correlate with teachers' years of experience. Furthermore, teachers did not accurately predict their students' responses to the CINS, and were surprised that Pre-International Baccalaureate Biology students had similar ideas to those in the general and ELL Biology courses.

Through the course of the study, drawing upon the Type 2 LP, the Daphne Project team encouraged the teachers to sort student responses, and to discuss sources of the ideas represented in them. In fact, during a planning time toward the beginning of the unit, two teachers in the study decided to sort student responses to the prompts, "Tell me everything you know about evolution."

Midway through the study, Teacher 3 explained that she had asked students to correct examples of student responses taken from the draft Type 2 LP. Then, the teacher categorized the student work into three groups, which she called "they go it", "they sort of got it", and "what the hell?" Upon exploring the contents of each group, the "they got it" responses were the correct answers, the "sort of got it" responses included student statements similar to, "The moths changed," but without explication of the mechanism was the student was applying for how the moths changed. Then the "what the hell?" pile was either misconceptions stated in the student ideas or just something that the teacher couldn't understand. This indicates that the teacher was not making a strong distinction between the types and sources of student alternative conceptions.

Post-interview results suggest that Teacher 3's perspective on categorizing student thinking was not unique. At the end of the study, each teacher was independently asked to complete a sorting task in which they were given seven different student responses from one of the formative assessments, and was asked to sort the responses into whatever categories made sense to them. Results of this interview task indicate that the teachers all agreed on two of the seven responses being 'correct,' and in general were able to group the ideas based on alternative conceptions together. However, the teachers fell into three groups with respect to the way they subdivided these categories (Figure 8); two of the teachers sorted the responses into four groups representing a continuum of student thinking; three of the teachers created three categories that included correct answers, incorrect answers, and unclear answers; and two of the teachers grouped the student responses into what they described as 'right' and 'wrong' answers.

For example, Teacher 1 (T1), explained that the responses represented "Kind of a continuum, as far as I'm concerned." In contrast, Teacher 7 (T7), explained that responses 3, 4, 6, and 7 were "Closer to being the right idea and the right concept," whereas responses 1, 2, and 5 were "the most way off."

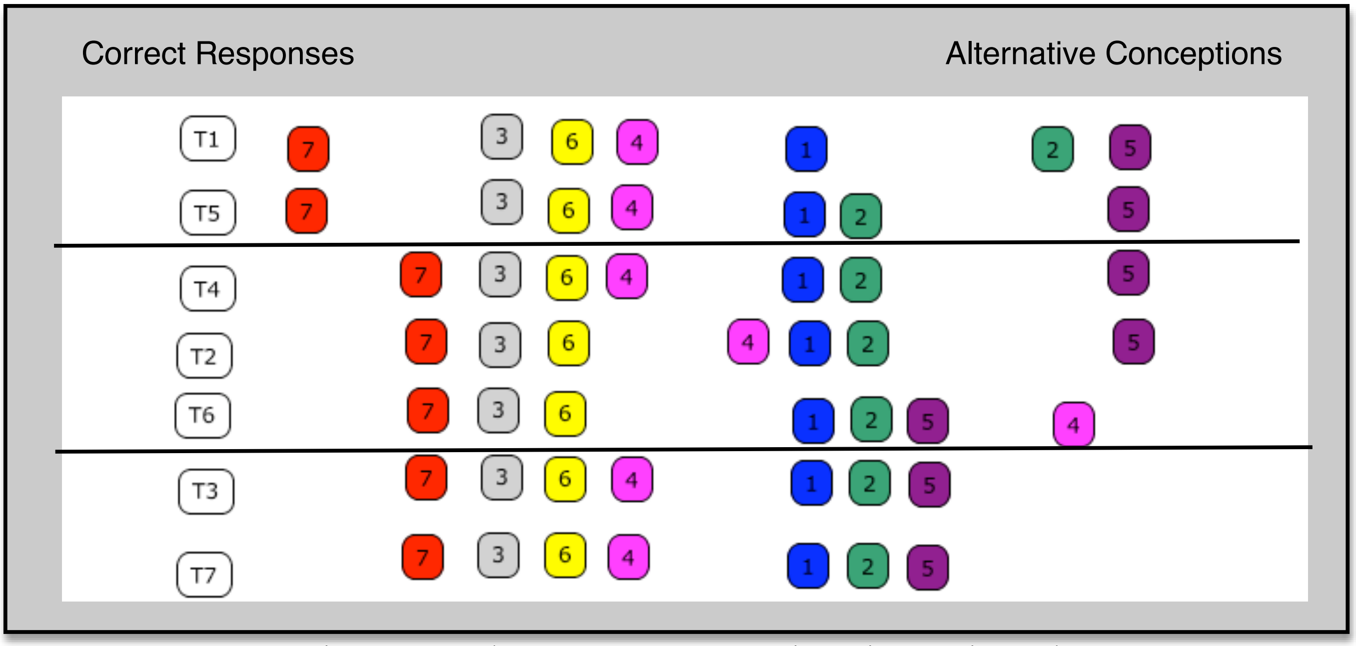

Figure 8. Teachers' responses to post-interview sorting task

These results indicate that, despite our efforts to explore student thinking with the teachers in association with the Type 2 LP, the teachers did not converge upon singular interpretations of the different pieces of student work. Furthermore, several of the teachers retained a correct/incorrect view of student responses. In future iterations of the project, we plan to speak more explicitly about the LP as a representation of ideas that goes beyond right and wrong, and the importance of identifying and acting upon different student responses.

Challenge 5. Anticipating feedback

Following criterion 4 stated above for LPs for teacher development, we focused during Y2 of the study on discussing and anticipating the feedback that teachers might give students to address particular ideas on the Type 2 LP. We attempted to determine what would go in the 'possible feedback' column of the LP by brainstorming as a project team, with the assistance of a biologist, as well as by speaking directly to the teachers during PLC meetings

These experiences indicate that thinking about feedback in advance is incredibly difficult, even with a group of 7 teachers and three researchers. In the end, while some ideas were tossed around the PLC, researchers in consultation with a biologist suggested most of the feedback ideas on the draft LP during Y2. Furthermore, when asked to anticipate what kinds of feedback they would give to the class that contained the seven responses teachers were asked to sort during the post-interview, most teachers described giving feedback to a whole class, rather than tailoring feedback to individual students.

While involving multiple stakeholders – teachers, researchers, and a biologist – proved effective in anticipating some feedback strategies, the ideas created were only hypothetical and have not all been road-tested. We are currently in the beginning stages of analyzing classroom videotapes made during Y2 to identify more feedback strategies that were successful.

For example, we collected videotape of Teacher 5 implementing the moth multiple-choice formative assessment, shown in Figure 4. In watching the video of that implementation, we identified both the ideas shared as they map onto the Type 2 LP for the origin of new traits, and the nature of the feedback provided by the teacher. Table 7 illustrates an example of this coding process.

Table 7. Sample transcript coded for LP ideas and feedback.

Transcript	*Notes*
Teacher 5: Alright, so, let's start with Emily. What compelled you about D? Why, what, made you think that was a good answer?	
Lindsay: Um, 'cause it talked about a mutation, and I just thought that the mutation made it more evident that it was from a change in the environment. We were saying that mutations [*inaudible*] but it was rather because they had adapted to their environment.	*Student Idea:* Need-Based Mutation
Teacher 5: OK. So when it said the moths themselves, which were able to make themselves a darker color to match the bark of the trees which have been darkened over time by industrial pollution. So you think that the moths changed color because of the pollution?	
Lindsay: Sure.	
Teacher 5: OK. Um, B.	
Student: That's exactly what I chose, I chose B.	
Teacher 5: Because B and D are actually very similar, right? OK. OK. So, do you guys remember, you guys remember this? So, it was this idea that in the US we have lots of resistant lice, right, to things like Rid-X, rid X? And the two ideas are, the two ideas are that in a given population of lice you have resistant versus non-resistant lice, or that exposure to the lice shampoo actually caused a mutation, right? And which one was it that we figured out?	*Feedback:* Referring students to evidence and conclusion from laboratory activity
Students: A.	
Teacher 5: Hypothesis A. OK. So let's go back to D. Option D. It says, the moths themselves, which were able to make themselves become a darker color to match the bark of the trees, basically resulted because of the pollution. Do you see any similarities there?	*Student Idea:* Environmentally-caused change *Feedback:* Teacher asks students to consider similarities between ideas
Student: Lamarck?	
Teacher 5: Exactly. So, this is hard, it's a really subtle difference, because for so long, you guys have learned that, you know, the	*Feedback:* Teacher tells students that

organisms change, adapt to fit their environment, and they do, but that's not a - that's the end result. Do you know what I mean? It's not the mechanism by how that happened.	organisms do not adapt to their environment

In this transcript excerpt, Teacher 5 calls on a student who selected one of the misconception-linked ideas from the moth formative assessment, and then takes three different approaches to provide feedback to that student – and others with similar ideas – about the shortcomings of her ideas. We are currently in the process of collecting these and other feedback instances from Y2 to add to the LP for Teacher Development.

Rather than push what has shown itself to be an unproductive strategy again in Y3 of the project, we intend to invite teachers to analyze classroom video with us by watching video of each other teaching the Natural Selection unit as a group, and discussing student thinking and feedback provided by teachers. We will take this new approach not only to add to the developing LP for TD, but also to help the teachers develop a better conception of the variety of strategies they might use to provide feedback to their students.

Challenge 6. Scaling Up

To date, the Daphne Project is two-thirds of the way through Phase 1, which has entailed development and piloting of the hypothesized LP for teacher development. In order to understand the potential impact of this LP on teacher knowledge and student learning, a separate Phase 2 research project will need to be completed with a new group of teachers, which raises a new set of challenges with respect to how this work can be 'scaled up' in a larger project.

The first concern raised deals with recruitment of teachers to participate in the project. The Daphne Project faced a slow start during Y2 as a result of difficulties in finding participant schools for the study. After a cumbersome review process with the partner school district that stretched on for nearly four months, principals were contacted by the school district to inform them of the potential for their teachers to participate in such a project. In all, 9 separate schools said 'no' at various levels for various reasons, including the district's choice to impose a new scope and sequence on the science departments already causing enough trouble for teachers, principals not wanting to overburden teachers when schools were being remodeled, coaches concerned that they would not have enough time to participate, and departments with several beginning teachers that thought they would be overwhelmed if they participated.

This experience calls into question the efficacy of scaling up a departmentally-based model of professional development around an LP, for the level to which the model can be scaled is necessarily limited by teachers' willingness to participate, no matter what their principals or districts might support.

A second challenge in scaling this type of research will involve the open question as to whether or not a LP will have utility beyond the teachers that developed it. While we were not concerned at the outset of the study, Teacher 2 raised the specter of the limited utility of the LP for Teacher Development in a post-interview. The teacher was describing how the Type 1 and Type 2 LP's might be combined:

Teacher 2: I was just trying to imagine what this would look like, and if you'd had something like, a specific, almost like a dichotomous key, like, you do this activity, you explain this, kids think this. Choose your own adventure, go to: Kids think A, go to this. And then go to a specific activity that puts that fire out. Now you've done that, you've done this group, maybe not target all the groups, not all the problems at once, maybe just one chunk of them. Okay, now you've done this activity, targeted that group, now do the exact same thing.

The interviewer then asked the teacher what he would do if a similar LP was developed for another unit by a group of teachers at another high school in the district, a question to which Teacher 2 responded in no uncertain terms:

Interviewer: If someone came to you…with something like that in the genetics unit –

Teacher 2: **I would throw it in the trash.**

Interviewer: Would you? Why?

Teacher 2: **Because it's not mine.** For some reason we like what we do, so if I made one it would be the best one ever, and if someone gave it to me, they're probably not very good at it. Even though that's the same thing I said to them, you can give them something, people still don't want to give you something, because **they want to feel like they're a part of it**. Yeah, I think process is what's most important.

This interaction caused the Daphne Project team to reflect upon the utility of LP's as tools to guide instruction, as well as to potentially help students learn. If researchers spend so much time in developing what we hope are the wave of future science education reform, what value do they have if teachers won't use them, and do the tools themselves retain any utility in isolation of the process by which they were developed?

While the previous questions are empirical, and can only be answered through further research, the Daphne Project hopes to address these questions during the next phases of research in two possible ways. First, we hope to involve teachers from the Phase 1 development as teacher leaders for the Phase 2 cohort so that teachers can directly address each others' concerns and speak directly to how the LP helped them in their own practice. The second way to address this challenge will be to build flexibility into the LP for TD at each school site, so that its contents incorporate the expertise of the local teachers as it relates to their own population of students. For example, this might involve developing formative assessments at each research site

Implications for 'Using' LPs for Teacher Development

A potential new 'use' for LPs is as tools that represent the knowledge necessary for teaching a particular concept, and therefore can help teachers to develop the PCK that will help them become more effective at impacting student learning. In order for LPs to be of use to the practicing teacher audience, they need to be created as tools that have instructional utility. Simply creating new versions of the standards that are sequenced progressions of student thinking is only a first step in the right direction; teachers also need a collection of activities and teaching strategies and activities to get students to share their thinking, guidance on how to recognize student thinking when it is shared, and concrete suggestions on how to provide effective feedback to students with a range of ideas. Furthermore, teachers need to have a sense of autonomy and ownership in any instructional tool they adopt.

The struggles represented by the teachers in the Daphne Project present a microcosm of the need for the educational research community to focus on its 'user' end when developing LPs. These challenges raise a number of questions for discussion by the LP research community. For example, what are the advantages and disadvantages of the two types of LPs to the teacher audience? What is the validity of simultaneously developing and using LPs at a school site? Is it possible to capture the relevant facets of PCK in a LP and, if so, how would a teacher not involved in the development of the LP use it as an instructional tool?

Although much interest is currently focused on the development of LPs for all avenues of science education, the community should proceed with caution. LPs, like all instructional tools, are relevant only in particular contexts, and cannot be treated as a universal solution to the challenges presented by science education reform. Much research is still needed to help determine the utility of LPs across multiple contexts and for a variety of user audiences. Given the research designs, criteria, and challenges presented in this paper, it is clear that this process of discovery will be slow at best, and the field is still in need of evidence to determine under what conditions, and for whom, LPs will be effective tools to guide curriculum and instruction.

References

Alonzo, A. C., & Steedle, J. T. (in press). Developing and Assessing a Force and Motion Learning Progression. *Science Education*.

American Association for the Advancement of Science. (1990). *Science for All Americans*. New York: Oxford University Press.

American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.

American Association for the Advancement of Science. (2001). *Atlas of Science LIteracy*. Washington, D.C.: National Academies Press.

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and Evaluation of the Conceptual Inventory of Natural Selection. *Journal of Research in Science Teaching, 39*(10), 952-978.

Boz, N., & Boz, Y. (2008). A Qualitative Case Study of Prospective Chemistry Teachers' Knowledge About Instructional Strategies: Introducing Particulate Theory. *Journal of Science Teacher Education, 19*, 135-156.

Briggs, D. (2008). *The Process of Developing and Validating Learning Progressions: Two Case Studies*. Paper presented at the American Educational Research Association.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic Assessment with Multiple-Choice Items. *Educational Assessment, 11*(1), 33-63.

Catley, K., Lehrer, R., & Reiser, B. (2005). *Tracing a Prospective Learning Progression for Developing Understanding of Evolution*: Paper Commissioned by the National Academies Committee on Test Design for K-12 Science Achievemento. Document Number)

Claesgens, J., Scalise, K., Wilson, M., & Stacy, A. (2009). Mapping Student Understanding in Chemistry: The Perspectives of Chemists. *Science Education, 93*(1), 56-85.

Collins, A. (1998). National Science Education Standards: A political document. *Journal of Research in Science Teaching, 35*(7), 711-727.

Duncan, R. G. (2008). A Learning Progression for Pedagogical Content Knowledge in Secondary Science Teaching: Project Summary [Electronic Version]. Retrieved March 30, 2009, from http://www.kstf.org/YoungScholarRavitDuncan.aspx

Duncan, R. G., Rogat, A., & Yarden, A. (in press). A learning progression for deepening students' understanding of modern genetics across the 5th-12th grades. . *Journal of Research in Science Teaching*.

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making Students' Thinking Explicit in Writing and Discussion: An Analysis of Formative Assessment Prompts. *Science Education, 92*, 799-824.

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P., Shavelson, R. J., et al. (2008). On the Fidelity of Implementing Embedded Formative Assessments and its Relation to Student Learning. *Applied Measurement in Education, 21*(4), 360-389.

Garvin-Doxas, K., & Klymkowsky, M. (2008). Understanding Randomness and Its Impact on Student Learning: Lessons Learned from Building the Biology Concept Inventory (BCI). *Life Sciences Education, 7*(2), 227-233.

Geddis, A. N. (1993). Transforming subject-matter knowledge: The role of pedagogical content knowledge in learning to reflect on teaching. *International Journal of Science Education, 15*, 673-683.
Gess-Newsome, J. (1999). Secondary Science Teachers' Knowledge and Beliefs about Subject Matter and Their Impact on Instruction. In J. Gess-Newsome & N. G. Lederman (Eds.), *Pedagogical Content Knowledge and Science Education* (pp. 51-94). The Netherlands: Kluwer Academic Publishers.
Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81-112.
Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*(3), 141-151.
Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the Ecological Assumption: The Relationship of Measure Schores to Classroom Teaching and Student Learning. *Measurement: Interdisciplinary Research & Perspective, 5*(2), 2007.
Kelemen, D., & Rosset, E. (2009). The Human Function Compunction: Teleological explanation in adults. *Cognition, 111*, 138-143.
Klymkowsky, M., Furtak, E. M., Cooper, M. M., Garvin-Doxas, K., & Gonzales, A. O. (draft). Molecular Obstacles to Understanding Evolution.
Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
Loughran, J., Mulhall, P., & Berry, A. (2004). In Search of Pedagogical Content Knowledge in Science: Developing Ways of Articulating and Documenting Professional Practice. *Journal of Research in Science Teaching, 41*(4), 370-391.
Mayr, E. (1997). *This is Biology*. Cambridge, MA: Harvard University Press.
Nagashima, S. O., Brown, N. J. S., Fu, A., Timms, M., & Wilson, M. (2008). *A Framework for Analyzing Reasoning in Written Assessments*. Paper presented at the Annual Meeting of the American Educational Research Associaton.
National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.
National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: National Academy Press.
National Research Council. (2001). *Classroom Assessment and the National Science Education Standards*. Washington, D.C.: National Academy Press.
National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8.* Washington, D.C.: National Academies Press.
Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington D.C.: National Academies Press.
Rudolph, J. L., & Stewart, J. (1998). Evolution and the Nature of Science: On the historical discord and its implications for science education. *Journal of Research in Science Teaching, 35*(10), 1069-1089.
Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal Formative Assessment and Scientific Inquiry: Exploring Teachers' Practices and Student Learning. *Educational Assessment, 11*(3&4), 237-263.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry. *Journal of Research in Science Teaching, 44*(1), 57-84.

Ruiz-Primo, M. A., Furtak, E. M., Ayala, C. C., Yin, Y., & Shavelson, R. J. (in press). Formative Assessment, Motivation, and Science Learning. In H. Andrade & G. Cizek (Eds.), *Handbook of Formative Assessment*. New York: Taylor & Francis.

Sadler, D. R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science, 18*, 119-144.

Sadler, P. M. (1998). Psychometric Models of Student Conceptions in Science; Reconciling Qualitative Studies and Distractor-Driven Assessment Instruments. *Journal of Research in Science Teaching, 35*(3), 265-296.

Shavelson, R. J., Yin, Y., Furtak, E. M., Ruiz-Primo, M. A., Ayala, C. C., Young, D. B., et al. (2008). On the Role and Impact of Formative Assessment on Science Inquiry Teaching and Learning. In J. Coffey, R. Douglas & C. Stearns (Eds.), *Assessing Science Learning* (pp. 21-36). Arlington, VA: NSTA Press.

Shepard, L. (2009). *From Advocacy to Adoption: Getting Formative Assessment into More Classrooms. Discussant Remarks*. Paper presented at the American Educational Research Association Annual Meeting.

Shtulman, A. (2006). Qualitative differences between naive and scientific theories of evolution. *Cognitive Psychology, 52*, 170-194.

Shulman, L. S. (1986). Those Who Understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Siegler, R. S. (1997). Concepts and Methods for Studying Cognitive Change. In E. Amsel & K. A. Renninger (Eds.), *Change and Development: Issues of Theory, Method, and Application* (pp. 77-97). Mahwah, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Crowley, K. (1991). The Microgenetic Method: A Direct Means for Studying Cognitive Development. *American Psychologist, 46*(6), 606-620.

Smith, M. L. (2006). Multiple Methodology in Education Research. In J. L. Green, G. Camilli & P. B. Elmore (Eds.), *Handbook of Complimentary Methods In Education Research* (pp. 457-475). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Talbot, R. M., Briggs, D. C., & Otero, V. K. (2009). *Can Science Teachers' Strategic Knowledge be Conceptualized as a Learning Progression?* Paper presented at the Annual Meeting of the American Educational Research Association.

Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, N.J.: Erlbaum.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P., Furtak, E. M., et al. (2008). On the Impact of Formative Assessment on Student Motivation, Achievement, and Conceptual Change. *Applied Measurement in Education, 21*(4), 335-359.

Table 6. Revised Type 2 LP for Teacher Development – Origin of New Traits

Name	Description	Assessment Prompts	Example Student Responses	Feedback Ideas
Random Mutations	Multiple kinds can occur (which can occur either spontaneously or in response to environmental mutagens), some of which accumulate over time and lead to new beneficial traits in organisms that are then selected by the environment. The effect of a particular mutation will depend upon the rest of the genome (and in fact the overall make of up of the population as a whole).	How do species change over Time?	A species changes over time because of random mutations and gene shuffling. Random mutations can cause a change in a species' gene pool. And gene shuffling is the different combinations of genes that come from the parents. If species are separated long enough, the species' gene pool changes.	
Environment causes change with genetic basis	Changes occur as a result of genetic mutations in direct response to the environment	Individual vs. Population Change	Animals mutate to fit in with their natural surroundings. So becoming darker helps to keep them in camouflage. If the bark gets darker the moth needs to go through some genetic mutations to match the bark.	Refer back to the different ways that changes in DNA can occur. Explicitly discuss the common misunderstanding that the environment causes advantageous 'mutants'.
Mutations Undefined	Student refers to mutations leading to new traits but does not describe a mechanism for how that happens.		If a mutation happens it can effect the whole species by creating a variety of differences from color change to more or less help against gathering food and protecting against predators. When the genes started to mutate the light color moths died leaving only the dark color	Explore framework for different kinds of mutations that go beyond those that are creative (including

Name	Description	Assessment Prompts	Example Student Responses	Feedback Ideas
			months to reproduce.	those that are neutral, deleterious, etc.).
Environment causes change/no genetic basis	Organisms change as a direct result of environmental changes; mechanism of transmission of changes is unclear. Eg. The environment changed the moths.	Moth Multiple-Choice plus Justification Correct the Misconceptions	Species change over time because of where they live. For example, a school of fish gets thrown on land because of a tsunami, the fish would more than likely die but one femail fish lays her eggs on land. When those fish are born, they will have no water around so they have to adapt to land.	Refer to in-class laboratory in which an example of natural selection was explored. Conduct simulation activity in which proportions of individuals in a population change over time as a result of environmental factors. Discuss the genetically-based ways that new traits can arise. Explore examples of student responses and identify alternative conceptions.
Anthropomorphic	Organisms consciously change themselves to adapt to environmental changes. Things that organisms learned in their lives are traits to be passed on.		Evolution is when an organism changes through time. For example cheetahs change by the way and by how fast they run. They learn to survive in their environment and to either be predators or prey.	
Eugenic	Organisms choose mates intentionally to create offspring that will be well suited to the environment		Organisms have to adapt to their surroundings, so if you live in a place where the holes and stuff are small, you will want to have a baby with someone small, and so eventually all the population will be small.	
No explicit mechanism	Mechanism for change in populations is unclear	All of the above	When the world is evolving and changing populations can't stay the same because they won't be able to survive. So this is why	Ask students to explain what they mean by ambiguous words such as 'adapt'

Name	Description	Assessment Prompts	Example Student Responses	Feedback Ideas
			species evolve as an environment changes species need to adapt to it.	and ‘change.’

Author Note

The research reported in this paper was supported by a Research Fellowship from the Knowles Science Teaching Foundation. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors. They do not necessarily represent the official views, opinions, or policy of the Knowles Science Teaching Foundation.

Some of the ideas presented in this paper appeared in an earlier form in Furtak, E. M. (2009). *Learning Progressions to Support Teacher Learning*. Paper presented at the Annual Meeting of the American Educational Research Association.

Acknowledgements

The author wishes to thank the teacher participants in the Daphne Project for their openness and support during this ongoing research, as well as the Daphne Project research team: Sarah Roberts, Deborah Morrison, Genevieve Turner, Shealyn Malone, and Prof. Michael Klymkowsky. In addition, Bud Talbot and Rick Duschl provided constructive feedback on earlier drafts of some sections of this paper.