

KNOWLEDGE TRANSFER ACROSS ASSESSMENT TECHNOLOGIES:
CONSIDERING THE ENTIRE RANGE OF FORMATIVE AND SUMMATIVE FUNCTIONS

by

STEVEN J. ZUIKER

(Under the Direction of Daniel T. Hickey)

ABSTRACT

Educational assessments reflect different, often contrasting, assumptions about learning and knowledge transfer. This study considers the value and function of assessment technologies that resonate with different general families of learning theory. Such consideration is conceptualized within a situated cognition framework in order to relate data across these technologies. Students in four high school science classes (n=52) enacted the *GenScope Assessment Project* curriculum on genetic inheritance and all assessment technologies. Pre-post administrations of two individual-oriented assessments generated significant quantitative gains, suggesting knowledge transfer in terms of their respective theoretical assumptions. A third set of assessment technologies generated event-based data about student participation, yielding qualitative profiles of one focal group's "knowledge practices." These profiles also compare the group's participation on the third set of technologies with their performance on the first two. This comparison begins to characterize the relative value and function of each technology within a multi-level, multi-type assessment framework.

INDEX WORDS: Transfer, Assessment, Discourse Analysis

KNOWLEDGE TRANSFER ACROSS ASSESSMENT TECHNOLOGIES:
CONSIDERING THE ENTIRE RANGE OF FORMATIVE AND SUMMATIVE FUNCTIONS

by

STEVEN J. ZUIKER

B. S., Psychology, University of Illinois at Urbana-Champaign, 1997

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment
of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2005

© 2005

Steven J. Zuiker

All Rights Reserved

KNOWLEDGE TRANSFER ACROSS ASSESSMENT TECHNOLOGIES:
CONSIDERING THE ENTIRE RANGE OF FORMATIVE AND SUMMATIVE FUNCTIONS

by

STEVEN J. ZUIKER

Major Professor: Daniel T. Hickey

Committee: Betsy Rymes
Paul Schutz

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2005

DEDICATION

“He who would learn to fly one day must first learn to stand and walk and run and climb and dance; one cannot fly into flying.”

-Friedrich Nietzsche

ACKNOWLEDGEMENTS

I extend inestimable gratitude to fellow annex members Marina Michael and Bryon Hand for their support. I also thank the many GenScope collaborators, but especially Nancy Schaefer and Art Russell for their effort and patience.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES.....	viii
LIST OF FIGURES	ix
CHAPTER	
1 INTRODUCTION.....	1
2 CONCEPTUAL FRAMEWORK.....	5
Transfer and Assessment.....	7
Formative Functions of Assessment	15
Scientific Argumentation.....	17
Discourse Analysis.....	19
3 METHOD	23
Participants	23
Design and Materials.....	24
Procedures.....	38
4 RESULTS	40
Class-Wide Individual Learning Outcome Measures	40
Pedigree Prediction Results of the Learning Outcome Measures.....	42
Discourse Analysis of Collective Participation in Pedigree Prediction Activities .	52
5 DISCUSSION	75

REFERENCES	81
APPENDICES	88
A Pedigree Prediction Dragon Investigation	88
B Unit Three Test Materials Related to Pedigree Prediction	90
C Final Exam Materials Related to Pedigree Prediction.....	93
D Test Review Steps	95
E NewWorm Performance Assessment	96
F Multiple Choice Test	106
G Transcription Key.....	109

LIST OF TABLES

	Page
Table 1: Focus Group Students' and Overall Gains Scores for Learning Outcome Measures	43
Table 2: Focus Group's Trajectories for Pedigree Chart Items on Learning Outcome Measures	51

LIST OF FIGURES

	Page
Figure 1: Pedigree window from the GenScope software environment	27
Figure 2: Pedigree chart from the unit three test	29
Figure 3: Sketch image from a videotaped activity	39
Figure 4: Items 14 and 15 from the near-transfer NewWorm performance assessment	45
Figure 5: Item 21 from near-transfer NewWorm performance assessment.....	47
Figure 6: Item 13 from far-transfer multiple choice test.....	48
Figure 7: Item 14 from far-transfer multiple choice test.....	49

CHAPTER 1

INTRODUCTION

The field of “educational assessment” encompasses a bewildering array of practices (National Research Council [NRC], 2001b). As this study illustrates, informal observations, quizzes, and standardized achievement tests each present and elicit different representations of experience. These differences, however, are not inherently bounded by theory. Rather, theories of knowing and learning provide backgrounds against which different representations of experience become interpretable, and sometimes relevant.

Applied to educational assessment, students and test developers (e.g., classroom teachers, book publishers, testing agencies) both do the work of representing curricular experiences. Assessment systems typically frame curricula according to a consistent set of theoretical assumptions. These assumptions underlie the kinds of representations of a curriculum that can be valued in assessment systems. Therefore assessment systems recognize representations only insofar as they resonate with or conform to the ways of knowing delineated by that system’s underlying theoretical assumptions. In these ways, the learning theories embodied in assessment systems commonly define the ways in which representations can be interpreted as knowledge.

For many innovative educational learning environments, the theories guiding what counts as knowing in a curriculum maintain tenuous relationships with the theories guiding what counts, or “what works” more broadly, in educational evaluation. Curriculum theorists often operate under different theoretical assumptions than assessment theorists. These discontinuities can misrepresent and de-value both curricula and assessments.

In the USA, the *No Child Left Behind Act*¹ identifies specific assessment practices (i.e., externally-developed, criterion-referenced achievement tests) as accountability levers for schooling and the primary instruments for driving educational reform. While these practices serve particular purposes, their diffusion into classroom assessment serves more peculiar purposes that arguably exceed their scope.

Federal guidelines increasingly treat achievement tests, when used in randomized trials, as the “gold standard” for evidence in educational research (NRC, 2002). However, these measures narrowly define what counts as knowing. Such narrow definitions may be necessary in the high stakes arenas of accountability-oriented educational reform and research funding, but they also inadvertently limit the potential for classroom assessment practice to impact measures developed from such definitions. At the same time, other classroom assessment practices (i.e., alternative assessments, performance assessments, portfolios, etc.) may be equally limiting when developed independent of the learning standards to which schools are held accountable. These limitations begin to illustrate functional differences in assessment practices and underscore the need to consider their inter-relationships.

This study breaks new ground in educational assessment by analyzing measures generally consistent with assumptions underlying three different families of theory about learning. Each view supports different assumptions about transfer with different implications for assessment. Bringing together data from all three viewpoints, the study considers the diversity of and discontinuity across assessment systems consistent with each view of learning.

This study specifically addresses how different assessment practices represent the ways that students know and learn about a 20-hour secondary life science curriculum featuring

¹ Detailed information is available at the following website: www.ed.gov/nclb.

innovative technology. The study emerged as part of a broader effort to improve classroom assessment and its alignment with external testing, and to investigate learning and transfer across perspectives.

The three assessment practices used in this study consider different data sets as evidence of learning introductory genetic inheritance. Using two quantitative data sets collected within quasi-experimental designs, the study first examines the following two hypotheses:

1. Learning about genetic inheritance in the GenScope curriculum transfers broadly to individual performance on conventional assessment practices.
2. Learning about a specific genetic inheritance topic (i.e., making predictions about a pedigree chart) transfers to individual performance on conventional assessment practices related to that topic.

These two hypotheses explore one of the central arguments in recent considerations of assessment and testing. A National Research Council expert panel published the report *Knowing What Students Know* (2001), which called for a more careful alignment of assessment with instruction based on what experts know about the development of reasoning in particular academic domains. Building on their call, this study uses one conventional measure that scaffolds the development of reasoning about genetic inheritance and one traditionally designed achievement test. Each set of data frames the alignment of instruction and assessment according to different theoretical assumptions. In this way, they present representations of instruction and elicit representations of student understanding differently.

A broader understanding of assessment also entails better understanding of the relationship between individual participation in assessment activities, and collective participation in the curricular routines. These curricular routines should relate in large measure to the

countable learning recognized in the above assessment practices. As a third assessment practice then, the study considers videotaped enactments of formative assessment activities featured in the curriculum. Event-based data addresses a more interpretive question:

3. What forms of collective participation do students develop or refine through formative assessment activities?

Finally, considering the inter-relationships between all three assessments practices, the study considers a final question:

4. How does the assessment of collective participation in curricular routines relate to conventional assessments of individual performance?

By addressing transfer and assessment across three data sets, this study develops understanding of the relationship between learning environments and three different assessment practices. In so doing, it seeks to highlight an emerging role that formative assessment practices can play in mediating the complex relationships between learning and instruction (e.g., Black & Wiliam, 1998; NRC, 2001b), and classroom assessment and external testing (e.g., NRC, 2003; Wilson, 2004).

CHAPTER 2

CONCEPTUAL FRAMEWORK

This study applies contemporary theories of language and learning to practical problems in science education and formative assessment. In particular, it considers a data consistent with the assumptions of three “grand theories” of learning. As such, the study employs both nomothetic (i.e., oriented towards abstract, general principles) and ideographic (i.e., oriented towards unique, particular understandings) approaches, and draws on an eclectic body of prior research.

Fundamentally, this study embraces contemporary sociocultural theories of teaching and learning. These theories assume that knowledge is fundamentally situated in the social practices of communities. These ritualized forms of meaningful participation are understood as goal-directed activities mediated by tools like technology and signs like language. Therefore practices necessarily reside between community members and their shared tools (Vygotsky, 1978; Wertsch, 1991). In particular, this study embraces a “situative” view that is commonly associated with Greeno (e.g., Greeno and the Middle School Mathematics through Application Project Group, 1998). The situative view treats both individual behavior and individual cognition as “special cases” of participation in cultural activities. The domain-specific concepts and skills commonly featured in teaching and testing function as “artifacts” of students’ collective participation in broader sociocultural rituals. These artifacts re-emerge time and again as teachers and students, texts, and tests “enact” curricular intentions in school contexts. The various facts, concepts, terms, statistics, etc., associated with a domain represent socially defined practices that

ultimately define particular domains of knowledge. Collectively, they can be described as *knowledge practices*. In this sense, a knowledge practice is a broad collection of simple and complex tools and signs defined by their service to particular domains of knowledge and the communities that utilize them. A competent individual is one who participates successfully in the knowledge practices associated with a particular domain of expertise. Learning then is characterized as increasingly regular and successful participation in specific domain knowledge practices. As described by Greeno et al. (1998), learning occurs as individuals and the contexts in which those individuals participate become familiar with the “constraints” and “affordances” that support meaningful participation in those knowledge practices. Constraints and affordances generally refer to physical and social aspects of activity contexts that simultaneously bound and scaffold successful participation. Learning is manifested in the knowledge practices that emerge in interactions negotiated within the constraints and across the affordances of particular contexts.

The three distinct sets of assessment data considered in this study cohere through this framework. Event-based data attempts to document the ritualization of specific knowledge practices (i.e., making predictions from pedigree charts) in small group activities. Two paper-and-pencil learning outcome measures attempt to assess ritualized artifacts that relate to the domain of introductory genetic inheritance. The constraints and affordances between individuals and each of these three assessment contexts arrange for the (re)emergence of knowledge practices.

Through this conceptual lens, the first section below develops a framework for understanding transfer and its relationship to different assessment technologies. The next section reviews research on the formative functions of assessment as they operate in assessment systems. The last two sections then consider argumentation as a kind of discursive practice. The

penultimate section addresses the practical use of argumentation in science classrooms and its potential service to the formative functions of assessment. The final section considers a theoretical perspective on argumentation as a linguistic resource in evidential systems like genetic science.

Transfer and Assessment

In general, the idea that what students learn in school serves them as citizens assumes that education transfers beyond schools. Transfer therefore is central to education; it also underlies assessment. As suggested in the introduction to this study, a coherent study of assessment practices should consider the assumptions about learning and transfer that underlie these practices.

Transfer generally refers to something that is learned in one context and later used in another. It has been explored differently according to the particular theory of learning and knowing that frames it. However, these various explorations have been usefully, though perhaps not unproblematically (e.g., Zimmerman, 1993), grouped into three families of learning theory that will be referred to as the *empiricist*, *rationalist*, and *sociocultural* traditions (cf., Case, 1996; Greeno, Collins, & Resnick, 1996; Greeno & the Middle School Mathematics through Application Project Group, 1997).²

Both empiricist and rationalist families of theory generally assume that knowledge and learning can be located in internalized structures within the brain. They diverge, however, in their conceptualizations of the relationship between internal and external elements of knowledge.

² Case (1996) labels these views *empiricist*, *rationalist*, and *socio-historic* while Greeno, et. al. (1997) refers to them as *behaviorist*, *cognitive*, and *situative*.

As detailed below, these differences fundamentally shape the assumed processes through which an individual acquires knowledge. Meanwhile, the sociocultural family of theories generally emphasizes the interdependence of internal and external, assuming knowing and learning to be an undifferentiated whole stretched across the individual and the material and social environment. Collectively, these views lead to different ways of defining transfer. The subsections below characterize these differences with implications for how each family of theories views assessment technologies.

Empiricist Views

The view of knowing and learning that is widely labeled “empiricist” assumes that active external forces shape a passive mind (Case, 1996; Greeno, et. al., 1997). The behaviorist models associated with Skinner and later in the human information-processing models following the “cognitive revolution” in the 1960’s each embody the empiricist perspective. Cognitive structures reflect an initial environment through associations wrought of interaction. These associations reduce experience to stimulus-response connections; combined and recombined, they add up to constellations of durable associations that constitute knowledge. Knowledge transfers when similarities between the components of different task environments enable individuals to employ that knowledge elsewhere (e.g., Thorndike’s (1906) notion of identical elements). Therefore only specific knowledge transfers, and only when either the external conditions or the internal mental actions related to them invokes that knowledge. A new or different task, like educational assessments, represents a potential transfer context to the degree that elements of the initial and transfer environments overlap.

Most standardized assessment practices embrace empiricist assumptions about knowing and learning. The carefully worded stems and response alternatives of multiple-choice items, for

example, often represent knowledge as connections between discrete associations and the ability to distinguish between more and less correct associations. This process exemplifies the underlying assumption that complex knowledge can and should be broken into its component connections. Such an assessment practice can then address the universe of domain knowledge by probing a sample of these associations. An individual's constellations of associations can be meaningfully represented by her responses to a small array of items. Further, drawing randomly from large pools of such items can generate assessments without bias towards any particular curriculum. Overall performance on arrays of randomly selected items statistically approximates individual understanding in order to warrant claims about learning and transfer. In this sense, student performance represents both the knowledge an individual possesses and suggests its potential to transfer beyond the classroom, a basic premise of schooling.

In the present study, the first of two learning outcome measures is a multiple choice achievement test that requires discrimination between more correct and less correct associations drawn from the domain of introductory genetics. This measure represents an empiricist measure and reflects the "peculiar" kinds of assessment practices that students must navigate in order to succeed in most schooling contexts (e.g., end-of-course exams, graduation tests, college entrance exams). They are peculiar practices within this conceptual framework because they isolate an aspect of a knowledge practice and abstract it, by degree, from the immediacies and particulars of curricular activities. Test developers design and control representations of these knowledge practices and typically disregard specific curricula in order to reduce bias. Therefore, items ultimately render impoverished profiles of students' knowledge practices in and of the curriculum. At best, they approximate individual understanding of a formal domain. Students who perform better on multiple-choice assessments likely understand the curriculum it assesses.

It is also possible that those who perform worse understand it as well. Better performing students, however, do understand something that worse performing students do not; they attune to the key constraints and affordances in the items and the test design.

Rationalist Views

Rationalist views of knowledge transfer, like empiricist views, feature cognitive structures but with positions generally antithetical to empiricist views. Rather than an active environment shaping a passive mind, the rationalist assumes an active mind making sense of a passive environment (Lightfoot & Cox, 1997). Cognitive structures are not so much constrained by the environment as the environment is limited by the ways in which these structures organize experience. The brain is a unique, sense-making organ that actively constructs flexible conceptual understandings (Bransford & Schwartz, 1999). From these many particular understandings come a variety of kinds of transfer. Each corresponds to a particular cognitive structure, but metacognition is key for all of them (Mayer & Whittrock, 1996). Transfer occurs both generally and specifically insofar as a general metacognitive skill set enables an individual to select appropriate domain knowledge and general principles, and to monitor their application in a new problem space. Transfer then is not a unitary phenomenon concerned with the degree of similarity, as empiricist views assume, but rather a complex set of different transfer phenomena (e.g., Barnett & Ceci, 2002) orchestrated to solve problems. These fundamental differences in empiricist and rationalist perspectives have led to corresponding variation in assessment.

The constellations of multiple-choice items utilized in empiricist assessment practices constrain the assessment of higher order, top-down processing rationalist models. The static associations embodied by a stem and its three to five response alternatives restrict the knowledge representations considered important in rationalist views. Rather than the external design of a

structured environment, rationalists assume the internal design of structured cognition. In order to elicit the dynamic conceptual understandings of cognitive structures, rationalist forms of assessment typically involve the creation of a product or demonstration of a process.

From this perspective, assessment must examine a student's ability to employ conceptual understandings on larger, extended tasks. In contrast to the more random process that follows from empiricist perspectives, assessment generally calls for systematic consideration of the complex knowledge structures that define a domain (e.g., Mislevy, Steinberg, & Almond, 2002). One common rationalist assessment practice is performance assessments. Performance assessments present open-ended, problem-solving activities that require comprehensive solutions (Payne, 1997). They include such things as essays, science experiments, and portfolios but also written responses to open-ended problems. Their distinguishing characteristic is the elicitation of explanations, or rationale, for phenomenon. Performance assessments attempt to make visible the conceptual lens through which individuals see and understand a problem or domain. Assuming an active mind and passive environment, rationalist forms of assessment shift from active test writers probing associations to active test takers presenting solutions. Therefore in the present study, the second of two leaning outcome measures is a performance assessment. Like multiple-choice assessments, this particular performance assessment is also a peculiar assessment practice according to the study's conceptual framework. It isolates and restricts practices, again by degree, to individual, paper-and-pencil activities. These items are divorced from the collective activities and associated knowledge practices that they seek to assess. They are increasingly mediated representations of a specific learning environment (e.g., different language and images) and they are increasingly constrained interactions within the domain (e.g., asynchronous paper-and-pencil "discourse").

In general, both the empiricist and rationalist learning outcome measures underscore views of knowledge as internalized structures, but each emphasizes different ends of an assumed duality between mind and nature in order to account for knowing and learning. According to views of situated cognition, both render impoverished understanding of the knowledge practices.

Sociocultural Views

Socioculturalist views begin from an altogether different set of assumptions. Rather than locating knowledge in kinds of internalized structures, these views stretch knowledge across the individual and the social and cultural activities in which the individual participates. In this way, learning is distributed across cognitive structures and environmental elements. Experience remains an undifferentiated whole, rather than a division into an internal and external duality. Therefore primary consideration is accorded to the dialectical relationship between individuals and their physical and social environments. This relationship is necessarily recursive and mutually constitutive (Beach, 1999). Events are partially determined by recurring social practices and artifacts of a ritualized activity and partially determined by an individuals' ongoing participation in that activity (Cole, 1996; Vygotsky, 1978). Knowledge is therefore inextricably bound to these rituals and embodied in the events that enact such rituals. These general aspects of the sociocultural family of theories have significant implications for the notion of transfer.

Transfer encompasses the constraints and affordances that scaffold meaningful participation in learning activities and later sustain it in transfer activities (Greeno, Smith, & Moore, 1993). There must be some perceived degree of continuity across activities and individuals must attune to this invariance (Greeno et al., 1998). In considering the backgrounds against which rational conceptualizations or empirical associations are foregrounded,

sociocultural perspectives attempt to recognize the broader contexts that constitute invariance, and thereby transfer (Packer, 2001).

By accounting for context and invariance, the notion of transfer becomes conceptually troubling. It is not so much that knowledge transfers, *per se*, but that sufficient invariance remains for the re-emergence of knowledge. In this sense, transfer is epiphenomenal. It is without theoretical value in many socioculturalist views because knowledge does not reside in internal structures exclusively (Beach, 1999). These views do not assume that transfer does not exist empirically (Greeno, 1997); they assume that context situates learning and knowing to a degree that the notion of transfer alone yields an impoverished account.

Learning and transfer environments necessarily blend because performance and context remain inextricably bound to one another (Lave, 1980/1997). Transfer hinges on continuities and therefore addresses that which remains consistent, both physically and socially. It concerns practices that coordinate environmental and cognitive structures in order to support and sustain meaningful activity from one context to another.

Socioculturalist assessment practices consider the ways in which an individual student engages and advances collective participation in inquiry (Greeno et al., 1996). This focus values the events and interactions that constitute knowledge practices as much as their artifacts. The present study uses “feedback conversations” as a way of assessing collective participation in the development, maintenance, and refinement of knowledge practices. Across periodic feedback conversations, students engage various domain-specific practices. Videotapes of these events afford interpretations of the regularities with which groups of students engage and advance discussions towards recognizable knowledge practices of introductory genetics.

The third and fourth goals of this study consider knowledge practices that surface in everyday classroom activity and judge how they might support student performance on the learning outcome measures. The idea of transfer is crucial to this consideration. It is the theoretical notion underlying the learning outcome measures considered in the study. As pre-post measures, both assessments are kinds of “experiments” to detect learning transfer, but in another sense, these assessments leverage curricular practice. There is value in bridging these views.

By trying to understand an experiment as an actual experience in the lives of subjects, by focusing on how the circumstances it presents *differ* [sic] from those of routine situations, and by successfully predicting performance differences in the separate contexts, theories (rather than experimental results) can become general without automatically becoming invalid at the same time. (Lave, 1980/1997, p. 67)

The learning outcome measures may be isolated and removed from classroom activities to the extent that they serve as experiments. The associations and concepts foregrounded in empiricist and rationalist assessments present a new set of circumstances that transform knowledge practices. While these views of transfer often reveal if knowledge transfers, they typically fail to reveal how it does or does not (e.g., Lobato, 2003). By understanding how these assessments mediate transformations of knowledge practices, assessment systems may be refined to better align domain practices with the varieties of assessment practices subsequently encountered. They may also be refined to diminish the performance differences attributable to facility with the assessment practices (e.g., test wiseness (Payne, 1997)).

This study assumes that assessment practices can be viewed as both transfer experiments and everyday practices. Further, in framing the learning outcome measures as impoverished representations of what students know, this study proceeds to investigate a socioculturalist assessment of knowledge practices in order to better understand the relationship between the

everyday practices of a curriculum and its various assessments and the impoverished assessments employed to evaluate the curriculum.

Formative Functions of Assessment

The present study builds from the widely held assumption that assessment serves multiple purposes (Bell & Cowie, 2001b). Considering its different roles in science education, a National Research Council report (1999) entitled *The Assessment of Science Meets the Science of Assessment Standards* identified three main functions:

1. to monitor educational progress or improvement
2. to provide teachers and students with feedback
3. to drive changes in practice and policy through accountability.

The report characterizes these as *summative*, *formative*, and *accountability* assessment respectively, and notes that summative and accountability functions pervade schooling, while more formative functions remain underdeveloped.

The formative functions of assessment commonly operate as a subordinate feature of summative and accountability assessments. In a review of assessment and classroom learning, Black and Wiliam (1998) contend that only when assessment information is *actually* (i.e., directly) used in the service of learning and instruction can it be effective. Such use occurs at administrative and policy levels (i.e., selecting and promoting students, and evaluating schools) but less commonly at the classroom level (i.e., refining pedagogy and curricula, directly advancing student understanding). In many ways, classroom assessment information remains unused or used in ways that undermine learning (NRC, 2001a). It generally informs decisions on remediation and curricular revision after students complete the assessed unit. This practice foregrounds summative functions of classroom assessment for students and formative functions for teachers. Assessment information informs refinements to pedagogy or a curriculum from unit

to unit or course to course, but for students such practice is only formative to the extent that continuous summative assessment provides feedback from unit to unit. Even then, feedback is indirect and secondary. Assessment information directly serves teachers and administrators who receive feedback on pedagogy and curriculum but only indirectly students who do not receive meaningful feedback on their understandings.

More coordinated approaches to formative assessment provide nuanced feedback not only about learning experiences but also during them. Another NRC report, *Classroom Assessment and the National Science Education Standards* (2001a), echoing Black and Wiliam (1998), concludes that the most important aspect of classroom assessment is the formative functions it can offer directly to students, particularly when these assessments are used as feedback to advance learning on specific responses. This focus has been developed by British researchers (Assessment Reform Group, 1999; Black & Wiliam, 1998; Gipps, 1999; Sadler, 1989, 1998; Torrance, 1993; Torrance & Pryor, 1998; Turnstall & Gipps, 1996) and advanced by others in science education (Bell & Cowie, 2001a, 2001b; Cowie & Bell, 1999; Delandshere, 2002; Duschl & Gitomer, 1997).

Assessment, in general, advances formative functions when “assessment information is used, by the teacher and pupils, to modify their work in order to make it more effective” (Black, 1995, as cited in Bell & Cowie, 2001b, p. 11). In order to develop such a process of learning, the broader project engineered formative functions of assessment through innovative “feedback conversations” (Hickey, Kindfield, Horwitz, & Christie, 2003; Hickey, Wolfe, & Kindfield, 2000; Hickey & Zuiker, 2003). The designed feedback activity simultaneously scaffolds student discourse towards (a) reconsideration of experiences within the learning environment and (b)

more abstract representations of experience embodied by the artifacts of an established domain of knowledge.

Related to the discussion of transfer above, assessments that traditionally operate as so many kinds of transfer experiments are appropriated in the routines of feedback conversations. The formative functions inherent to these assessments scaffold student discourse. Importantly, the construction of assessment questions and the feedback test review steps situate feedback conversations within a scientific framework consistent with sociocultural forms of assessment.

Scientific Argumentation

Student participation is an essential part of the formative functions of assessment but also for sociocultural forms of evaluation (Greeno et al., 1996). In the present study, scientific argumentation shapes the central classroom assessment activity called feedback conversations. Through these activities, small groups evaluate and further develop domain practices. Several designed features of the curriculum structure a student-centered, collaborative approach to argumentation. This design seeks to meaningfully engage small groups of learners in assessing themselves and others, not just being assessed by the teacher. It also involves learners in the process of developing and maintaining standards and judging understandings.

Argumentation has been studied widely since Toulmin's seminal work, *The Uses of Argument* (1958/2003), formally distinguished it from logic. Later, a series of science studies implicated social processes like argumentation as central to establishing scientific knowledge (Latour, 1987; Latour & Woolgar, 1986). Science education researchers who hold more sociocultural orientations, in turn, now argue that in order for learning environments to reflect the socially constructed nature of scientific knowledge, they must give higher priority to discursive practices, especially argumentation (Driver, Newton, & Osborne, 2000; Kelly &

Duschl, 2002). To this end, a process of argumentation can be generally described as the formulation of knowledge claims and warrants to support those claims, the evaluation of warrants, and subsequent negotiation of both based on evidence (Driver et al., 2000). The discursive practices associated with formulating claims, warrants, and counters to each are increasingly recognized as a valuable and realistic process for inquiring about scientific content. Further, discursive practice in general and argumentation as a particular manifestation in science are both key to knowing and learning about scientific domains (Harré & Gillett, 1994; Lemke, 1990; Wertsch, 1991).

Argumentation situates ways of knowing in authentic, scientific practice. Students experience curricular content through a process of scientific argumentation. In so doing, it provides opportunities to learn not only science, but about science. These facets of scientific content and process are always already blended. As Lave & Wenger (1991) maintain, “the purpose [of discursive practice] is not to learn *from* talk as a substitute for legitimate peripheral participation; it is to learn *to* talk as a key to legitimate peripheral participation” (p. 109). Engaging in discursive practice is a process of learning in that it is a process of increasingly rich participation in the practices of the community of scientists. For example, Roth (1996) illustrates how high school students’ appropriation of the language of physics is bound to knowing about physics. In the broader GenScope project, scientific argumentation scaffolds the formative functions of feedback conversations, situating knowledge practices within authentic scientific practice. It structures a learning culture by re-engaging and re-presenting curricular activities and assessment items in small group, student-centered conversations. Together, these aspects of scientific argumentation engender data valued in sociocultural theories.

Argumentation generates information about how participants talk about prior learning. It also affords students opportunities to transform their present knowing. Further, as a structured activity in classrooms, the argumentation fostered in feedback conversations serves as event-based data for sociocultural considerations of domain knowledge practices. Student discourse relates to their written representations on performance assessment questions and multiple-choice items, providing information about the transfer of knowledge.

Discourse Analysis

Beyond domain-specific practices, the present study also considers participation in broader discursive practices and the ways they organize conversations during curricular activities, as well as the values and limitations of these practices. Discourse analysis is one method that considers the use of language as a tool for constructing knowledge and social structures in educational settings (Adger, 2003).

Language is multi-layered and multi-faceted. As an exterior layer, the lexicon and grammar of a language characterize a state of affairs. Differences in word choice and order can have far reaching implications in meaning. This is especially true in science, given the precision of scientific explanations (Lemke, 1990). As a more interior, pragmatic layer, language can be viewed as action, as doing things. The idea of language-as-action incorporates the layers of lexical and grammatical meaning in order to consider its functional meaning. In this sense, all language is dynamic, open, and contingent (Bakhtin, 1998). Relevant to this study, discourse coordinates language and gesture towards the interactional achievement of certain actions (and not others). Further, in scientific activities a variety of additional resources, like charts, statistics, and diagrams also serve meaning making (Lemke, 1990; Ochs, Gonzales, & Jacoby, 1996). Meaning then is not in lexical, grammatical, or syntactic functions alone but necessarily the

context of their use as well. Rather than distinguishing a text from its context, these two aspects remain together as intertexts of a whole (Lemke, 1995). In this way, the study of discursive practices, and educational assessment practices therein, is also the study of social interaction.

Social interaction can be considered from a variety of viewpoints. Heller (2003) broadly contrasts perspectives as resembling either “discourse analysis of interaction” or “interactional analysis of discourse.” In the former, discourse is generally considered to be the observed, interactional mechanics at local and universal levels and can be categorized as *ethnomethodological* (e.g., Garfinkel, 1967; Sacks, 1992a, 1992b). Analysis focuses on the “orderly character of everyday life” (Garfinkel & Rawls, 2002, p. 5) grounded in local conversational norms. In the latter interactional analysis, discourse transcends local interaction and can be categorized as *interpretive* (e.g., Gumperz, 1982; Halliday, 1978). Analysis focuses on the interactional achievements of different forms of discourse. Ultimately however, whether discourse analysis is “of interaction” and thereby rooted in specific events (e.g., greetings) or “interactional” and therefore extended to transcendent frames (e.g., schooling) depends on the theoretical orientation of the analysis.

In an effort to understand formative assessment, this study employs an interpretive perspective. Through the interactional analysis of discourse, it considers small group enactments across various curricular activities. Such a functional approach seeks to consider the immediacies of empirical, discursive interaction with the intermediacies of local design (e.g., curricula or lessons) and transcendent influence (e.g., schools, neighborhood, family).

Language socialization is one interpretive lens for analyzing classroom interaction in this way. According to Ochs (1996), *language socialization* assumes “that language practices are socially organized and that, as novices recurrently engage in these practices with more expert

members of society, they develop an understanding of social actions, events, emotions, esthetics, knowledgeability, statuses, relationships, and other socio-cultural phenomena” (p. 408).

Dimensions of social organization include social identity, social acts and activities, and both affective and epistemic stances. Together, these dimensions coordinate meaning making. They are dynamic and mediated such that socialization is always a constant synthesis of the various meanings achieved through interaction.

Interactions in classroom environments operate predominantly through language but also other semiotic mediators. These mediators include texts (e.g., activity sheets, assessments, tests, text books), gesture (Roth, 2001, 2003), the classroom space, and material objects or representations. Science education contexts also incorporate a variety of additional mediational means (Lemke, 1990) like visual representations (e.g., charts, drawings, tables), mathematical symbolism (e.g., equations and quantitative relationships such as ratios, percentages, and probabilities), and experimental operations (e.g., employing Punnett squares in genetics), as well as scientific “Discourse” as not just scientific talk but a way of talking and valuing science (Gee, 1999, 2003a, 2003b), all of which underscore science as an evidential system.

Lexical and grammatical systems like scientific discourse can be pragmatically analyzed by foregrounding the more complex and interesting situational dimensions. In the case of evidential systems more broadly, Ochs (1996) speculates that intuitive links between evidence and epistemology foreground *epistemic stance* in the service of evidential forms of language. She characterizes epistemic stance as utterances that addresses the certainty (or doubt), reliability, or limitations of a proposition (cf. claims), including comments on the source of information (cf. warrants). Chafe and Nichols (1986), for example, considered evidentiality (a term loosely interchangeable with epistemic stance) as a dimension of social interaction across cultures. They

detailed seven “modes of knowing”: belief, induction, deduction, sensory evidence, hearsay evidence, hedges, and expectations. Each of these modes foregrounds a kind of epistemic stance that operates in evidential systems like scientific argumentation.

CHAPTER 3

METHOD

Assessing and comparing student learning is a complex problem. The present study addresses this challenge by collecting data consistent with assumptions for each of the three “grand theories” presented above. It was conducted in part to break new ground in assessment research methodology. Traditionally, assessment scores have been treated as objective evidence of learning and knowledge transfer. Contemporary situative views of learning are increasingly being applied to assessment research, leading some theorists to advance an alternative characterization of assessment practices and what the resulting scores mean (Delandshere, 2002; Gipps, 1999; Hickey & Zuiker, 2003).

Participants

The study took place in a suburban high school in the southeast United States. The school serves an almost entirely African-American student population (99.5%) from lower-SES communities in which roughly 30% of students qualified for the federal lunch subsidy. The school typically posted school-wide achievement scores that were below average but higher than most other schools in the district that also served predominantly African American students. Published figures reported that 61% of these students passed the science component of the high school graduation test on their first try. In particular, this study considers students in four life science classes taught by Mr. N, who was recruited via a request distributed to his district science coordinator. For his participation, he received a \$600 honorarium for each of the three years. Mr. N has a bachelor’s degree in biology and had been at the school for five years. Further, he had

participated in GenScope implementations in each of the two previous school years. Mr. N. therefore was familiar with the software and the assessment model as well as the ways in which students engaged both. He worked closely with the research team and played a central role in refining the curriculum.

The GenScope curriculum served as regular instruction therefore all students participated in all activities. 57 students agreed to complete assessments of introductory genetic inheritance content knowledge before and after the curriculum. 6 students in each class further agreed to be videotaped while participating in select small group activities. The teacher grouped the 6 students in each class into two triads. This subset of video groups, totaling 24 students, is a convenience sample selected on the basis of good attendance. One triad was later selected for detailed micro-analysis.

The selection of a triad for analysis began with a preliminary review of all eight videotaped groups. One group was chosen as a paradigmatic case (Flyvbjerg, 2001) based on the quality of interactions between group members and engagement in curricular activities. The group included two African-American boys named Toby and Brian (all names are pseudonyms). The third member of the group withdrew assent to be videotaped before the activities in this analysis occurred. Both Toby and Brian participated in the GenScope curriculum during Mr. N's final class each day.

Design and Materials

This study considers a 20-hour curriculum on introductory genetic inheritance, including classroom assessments, and three methods for understanding student learning. The design and materials reflect the larger project in which this study is situated. The broader objectives of this project included iterative refinements to formative classroom assessment practices, and

examination of related issues concerning student motivation (Hickey, 2003). Essentially, the research team implemented successive designs of formative classroom assessment practices to directly improve (1) classroom discourse and scientific argumentation and to do so in ways that indirectly impact performance on both (2) a conventional performance assessment and (3) a conventional multiple-choice achievement test.

Curriculum

As elaborated below, the curriculum structures twenty class periods into three thematic units. Each unit lasts five to six class periods, beginning with a series of investigations that use the GenScope software then ending with a classroom assessment and formative feedback activity. The curriculum concludes with a shorter, three-period final exam unit. During these periods, students review previous unit assessments during one period, complete the final exam during the next, and conclude with a final feedback conversation.

GenScope Software

Genetic processes occur along scales too small or slow to observe directly, yet often represent secondary school students' first formal exposure to probabilistic reasoning (Stewart & Hafner, 1994). These factors create fundamental challenges addressed by the GenScope software. GenScope is an open-ended software tool that creates "windows" for viewing genetics processes. Each of six windows represents concepts and processes at a different level of biological organization. Students can manipulate features displayed in any of the windows then observe the implications of this change as it cascades up and down levels. This synchrony enables students to view genetics at work across scales. At the same time, manipulations illustrate the natural variation that is a part of genetics. Outcomes remain within general proportion to the underlying probabilities of that manipulation, but are never exactly the same.

By randomizing the specific results of any manipulation, the software represents the probabilistic operations that characterize genetic processes. Therefore, along small scales (i.e., cellular processes like meiosis and mitosis) and slow scales (i.e., intergenerational pedigree and population dynamics), GenScope makes visible the elusive processes that define inheritance.

These observations within and across levels also provide data and evidence that enable scientific inquiry and argumentation. In this way, GenScope scaffolds engagement in the actual practices through which geneticists solve problems, but the software alone is only a medium for scientific inquiry.

Dragon Investigations

Each of three unit cycles includes two to three investigations across three to four class periods. The term *Dragon Investigation* describes the set of open-response questions that scaffold knowledge practices associated with a unit. Each Dragon Investigation includes a subset of questions for a whole class activity and one for a small group activity. In the whole class activity, the teacher models inquiry using GenScope then facilitates a whole class discussion to answer the questions. In small group activity, student triads work collaboratively to manipulate features in GenScope windows then use the generated evidence to answer the questions. The teacher concludes each Dragon Investigation with a whole class review. Finally, these investigations are reconsidered during classroom assessments and formative feedback activities. This study considers the majority of these investigations only indirectly, through individual performance on the learning outcome measures detailed below. However, one Dragon Investigation is considered more closely.

This study focuses on the knowledge practices associated with predicting pedigrees using charts. These practices are featured in the Dragon Investigation titled *Pedigree Prediction* (see

Appendix A). The Pedigree Prediction Dragon Investigation targets a particularly challenging aspect of introductory inheritance because learners must interpret complex patterns that determine the way inheritance operates for a given characteristic. Making these predictions requires that students solve the kinds of problems that define expertise in the domain of genetics. In order to arrive at a solution, the learner must reason backward. They construct the patterns in a pedigree chart in order to predict the genotypes that effected that pattern. Only by integrating three interrelated knowledge practices from prior Dragon Investigations can a student begin to successfully predict all genotypes. These three practices are termed *dominance relationships*, *allele relationships*, and *chromosome types*. In a sense, Pedigree Prediction represents the confluence of these three knowledge practices to develop a fourth, termed *predicting genotypes*. Each of these practices will be explicated in the analyses below, but hinge on a fundamental understanding of the charts themselves.

The Pedigree Prediction investigation features multifaceted pedigree charts that present a distribution of traits across generations for a given characteristic. Figure 1 below is a screenshot of a pedigree chart that appears in the GenScope pedigree window.

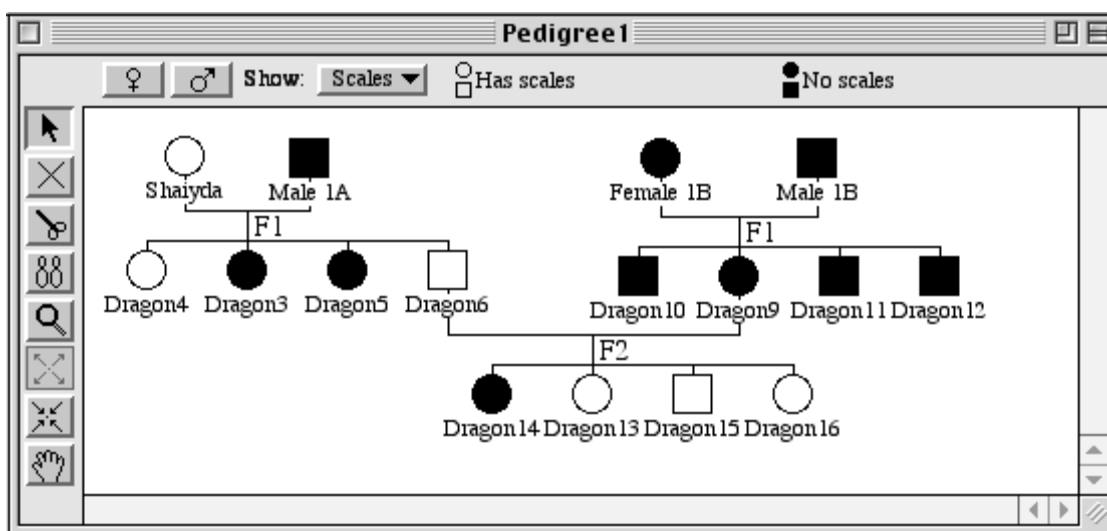


Figure 1. Pedigree chart in the GenScope pedigree window

The chart in this window shows the distribution of two traits – “has scales” and “no scales” - for the characteristic scales across three generations. Learners must interpret shape- and shading-coded symbols in terms of their lined arrangements in order to discern the pattern embedded therein.

Classroom Assessment and Formative Feedback Activities

Reflecting the goal of the broader research project, classroom assessment activities occupied nearly half of the 20-hour curriculum. Such an intensive assessment model can be understood as the range of forms and procedures that generate information about learning, termed “assessment technologies,” and also as the enactment of these technologies by students, teachers, and researchers, termed an “assessment system” (Hall, Knudsen, & Greeno, 1996). Whereas assessment technologies commonly serve assessment systems that operate apart from learning and instruction, this formative assessment system attempts to integrate its technologies with learning and instruction. Therefore the system seeks to employ technologies to formative ends.

Assessment technologies. There are three different kinds of activities in which assessment technologies operate: unit tests, the final exam, and pre-post achievement measures. Each thematic unit includes a unit test; the curriculum includes a cumulative final exam; and, beyond the curriculum, the research study adds pre-post achievement measures. The three unit tests and final exam integrate assessment activities with learning and instruction while the pre-post assessment activities operate independently. For this reason, the unit tests and final exam assessment activities coordinate four assessment technologies.

The first assessment technology is the actual unit tests and exam. Both are performance assessments that attempt to make visible students' conceptual understandings and are generally consistent with the rationalist family of theories. A unit test includes multiple sections, one for each Dragon Investigation in that particular unit. Each section has between three and eight questions related to a comprehensive scenario that features organisms identical to the GenScope environment and similar traits and pedigree charts. The final exam, in turn, presents twenty-nine questions about five scenarios, which incorporate all aspects of the curriculum. Concerning the Pedigree Prediction, both technologies scaffold predictions by dividing the interpretation of pedigree charts into the four practices mentioned above (e.g., dominance relationships) (see Appendices B and C for unit three test and final exam items respectively). Figure 2 below is a pedigree chart featured in the unit three test.

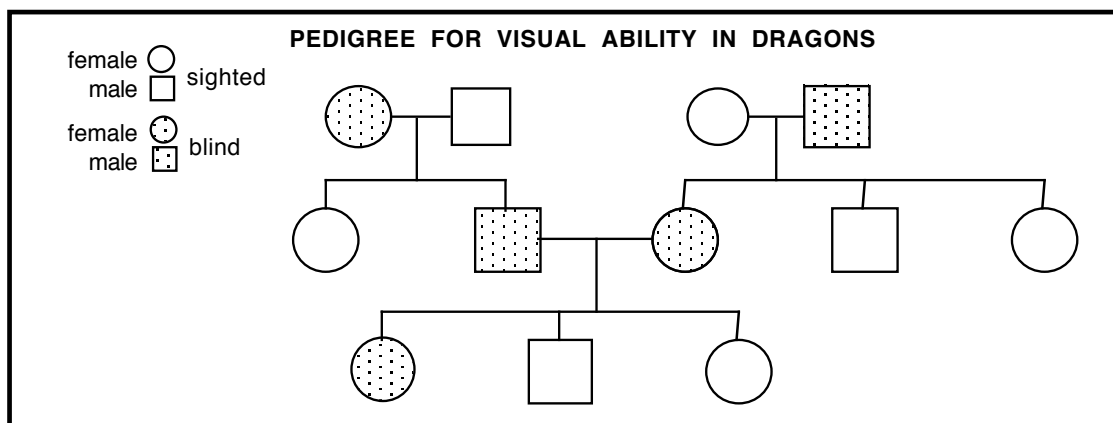


Figure 2. Pedigree chart from the unit three test

This chart exhibits subtle differences in representation relative to the GenScope pedigree window shown above. It employs the same basic geometric, shading, and line conventions without the GenScope toolbars or individual organism labels. The diagram title and key are positioned

differently as well. These differences begin to underscore the contrasting representations of learning experiences that assessment items present and elicit.

The second assessment technology, called *Answer Explanations*, is a set of rubrics oriented towards students. As such, these rubrics provide feedback to the learner rather than rules for a scorer. Answer Explanations are paragraph-long accounts of the reasoning behind each unit test and final exam item's solution (see Appendices B and C for unit three test and final exam Answer Explanations respectively). Answer Explanations often extend beyond the genetics content necessary to solve a problem and are written in relatively dense scientific prose. The length and genre serve to scaffold a triad's discourse around the domain; at the same time, they do not directly state answers.

The third assessment technology is *Judge Your Understanding* rubrics. Like Answer Explanations, they are also learner-oriented. They overview the domain practices presented in a unit test (but not the final exam) and identify which questions relate to each practice (see Appendix B for unit three test rubric). Students evaluate whether or not they understand each practice then rate their overall understanding on a four-point scale as either beginning, developing, accomplished, or exemplary. Students complete Judge Your Understanding rubrics only after completing a unit test and reviewing the associated Answer Explanations. The use of these three forms is coordinated by a step-by-step procedure.

The fourth and final technology is a four-step, formative process called the *Test Review Steps*. These steps coordinate the use of the above forms for a particular assessment and feedback activity. (See Appendix B for the text of the Test review steps.) For each individual item, each student states an answer and reasons for that answer (steps 1). The group then negotiates a "best solution" (step 2) and reads aloud the item's answer explanation (step 3). Through further

collective discussion, the group then either reaches consensus about “why the correct answer is correct and wrong answers are wrong” or agrees to disagree (step 4). By following this routine, groups ideally develop a sense of scientific argumentation by making and warranting knowledge claims in steps 1 and 2, then defending or refining claims in light of evidence in steps 3 and 4.

Together, these technologies (re)engage students with knowledge featured in Dragon Investigations, and therefore presents the possibility of knowledge transfer. At the same time, each of these technologies mediates differently the (re)presentation of knowledge and therefore can likewise be viewed as a transformation of knowing. The next section therefore details the interrelationships between unit tests and the final exam.

Assessment system. The enactment of the above four technologies by students, teachers, and researchers can be understood as an “assessment system” (Hall et al., 1996). The formative model operating in The GenScope Assessment Project attempts to design and refine the assessment technologies above towards a system in which they are increasingly used “by the teacher and pupils, to modify their work in order to make it more effective” (Black, 1995, as cited in Bell & Cowie, 2001b, p. 11). Through iterative design, the project sought to promote, sustain, and better understand the assessment systems that emerged across different enactments. The overview below details the general features of the system as designed.

The assessment system begins with administration of unit tests. Items promote central concepts and probe central understandings operating in Dragon Investigations, beginning a feedback process in that it represents only key features of students’ curricular experiences. Students complete unit tests and all assessments individually. The teacher collects and reviews but does not grade or otherwise evaluate students’ responses. Administering unit tests in this way maintains formal goals for student learning. Not assigning marks or grades, meanwhile, keeps

these formal goals from undermining the formative functions of unit tests. This balancing of functions makes these semi-formal assessments. They are designed to support individual accountability through formal administration and an active, ongoing culture of learning by means of an informal evaluation presented below in the second part of the assessment system. In the case of the final exam however, the teacher grades student responses and writes down a percentage score and letter grade. This re-invokes summative assessment practices and, for the purposes of the research project, provides opportunities to consider the sustainability of the formative assessment practices that students develop across the semi-formal unit tests. Therefore, administration, particularly the teacher's formal role as proctor and later as evaluator, begins to coordinate the assessment technologies.

In the second part of the assessment system, triads review assessments and feedback materials as part of a “feedback conversation.” Feedback conversations are collaborative evaluations that provide students with opportunities to discuss their responses and to consider additional information that further advances knowledge practices. These activities are supported by two other assessment technologies. The Test Review Steps orchestrate discourse around the particular assessment, students' written responses, and the particular Answer Explanations. The Answer Explanations, meanwhile, scaffold discourse towards central understandings featured in the assessment itself.

This two-part process of formal administration and informal evaluation repeats across three unit test cycles. Such recursion affords increasing engagement in authentic practices associated with scientific argumentation during the immediate feedback conversation. Repeatedly cycling through this process also affords feedback on students' experiences during the unit, specifically on knowledge practices featured in Dragon Investigations. Further, to the

extent that recursion of the assessment system ritualizes argumentation practices, the system develops a kind of forward feedback on the curriculum's inquiry process. Reflecting on evaluations of knowledge practices from previous investigations begins to frame the utility of future Dragon Investigations; it begins to socialize students into scientific inquiry and argumentation. At the end of the unit test cycles, the final exam serves to assess the curriculum through a similar two-part process.

The exam cycle includes three activities, beginning with a whole class review. During this class period, the teacher encourages students to review their unit tests, explaining that unit test items parallel final exam items and therefore understanding unit tests is key to doing well on the final exam. Students complete the exam on the next day then participate in a final feedback conversation with graded exams.

The GenScope Assessment Project's formative model operates on several levels. First, feedback can transform knowledge practices towards established domain understandings. Second, formative feedback can transform future interactions with GenScope and Dragon Investigation by establishing how prior interactions served them in solving problems on the assessment and defending them during the feedback conversation. In sum, unit tests coordinate a systematic and formative re-exploration of each unit cycle while the final exam provides a cumulative re-exploration across cycles.

Together, these four technologies shape an assessment system that is both informative to students and the teacher, and that advances a process of learning. It also provides useful information for considering how students use what they learned in any given Dragon Investigation across several new situations, which is the focus of this study.

Learning Outcome Measures

Three learning outcome measures provide three different representations of the domain. Each measure uniquely portrays one of the three views of transfer and assessment detailed above. Each measure also relates to the curriculum with differing degrees of instructional sensitivity. Similarities between curricular activities and either assessment items or student discourse characterize their closeness. Instructional sensitivity, in this sense, can be framed along a continuum from near to far, and thereby near-transfer and far-transfer measures of learning.

Near-Transfer NewWorm Performance Assessment

It is an abbreviated version of a paper-and-pencil performance assessment of introductory genetics called the *NewWorm* (see Appendix E). The *NewWorm* was developed in previous research with GenScope and is detailed in Hickey, Kindfield, Horwitz, & Christie (2003). Briefly summarized, the assessment is 25 open-response items involving a fanciful worm species whose genetics are similar to GenScope dragons. The ordering of these items follows a developmental model of expertise that scaffolds performance across increasingly complex problems and accurately assesses a broad range of expertise (Kindfield, Hickey, & Yessis, 1999).

The *NewWorm* reflects the rationalist perspective on transfer and assessment. It assesses understanding of the higher-level concepts that students presumably constructed to make sense of and complete curricular activities. Students employ these conceptual understandings to solve *NewWorm* problems that stem from domain concepts such as predicting pedigrees.

The *NewWorm* is relatively sensitive to instruction across the curriculum. *NewWorm* items resemble GenScope unit test and final exam problems. The items assess understanding of the domain only insofar as it was addressed in the curriculum. Therefore *NewWorm* does not address some common topics included in many textbooks, such as Mendelian genetics. Students

develop a degree of familiarity with the kinds of items they answer in the NewWorm. In turn, NewWorm items only assess familiar aspects of the domain. For these reasons, the NewWorm resonates closely with the GenScope curriculum and maintains a degree of instructional sensitivity that warrants its function as a near-transfer measure of individual understanding.

Far-Transfer Multiple-Choice Test

The far-transfer multiple-choice test (multiple-choice test) includes 13 multiple-choice items from two pools of items (see Appendix F). Five items were randomly selected from a pool of 45 released SAT II biology content test items that related to genetics. The remaining 9 items were randomly selected from 39 items in the genetics chapter test bank of a textbook's teacher supplement.

The multiple-choice test reflects the empiricist perspective on transfer and assessment. The stems and response alternatives in each item present discrete associations that assess components of complex knowledge. The array of 13 items, together, approximates individual understanding and renders measures of learning and transfer.

These items also approximate differences between groups of individuals. The random selection of test items reduces the likelihood that the test is biased towards any one curriculum, enabling its use as a proxy for performance on high-stakes tests. It follows that because the assessment is relatively impartial, it is also relatively less sensitive to instruction, warranting its use as a far-transfer measure.

The analysis of both near- and far-transfer learning measures employs multifaceted Rasch scaling by means of Facets software (Linacre, 1989). This item response theory procedure models a single linear scale on which both the relative difficulty of items and the relative proficiency of individual students lie. Items answered correctly by a small percentage of students

contribute more to individual proficiency while those answered correctly by only proficient students contribute more to item difficulty. Scaling in this way equates differences between items and students independent of the location of either one on the scale. Therefore, scaling generates a wider, more accurate estimate of difficulty and performance than simply summing correct responses. It likewise affords direct comparisons of items and students. The procedure also estimates the precision and reliability of the model, and estimates the degree to which each individual's pattern across items and each item's pattern across students fit the model's prediction. All results appear on a T-scale with a mean of 50 and standard deviation of 10.

Discourse Analysis of Small Group Activities

The analytical goal of this aspect of the study is to develop an understanding of emergent knowledge practices across the Pedigree Prediction Dragon Investigation, Unit Three Feedback Conversation, and Final Exam Feedback Conversation. These activities together coordinate the formative functions of the assessment system and present a considerable corpus of data with which to explore collective participation in learning and knowing about genetic inheritance. As discussed in the Discourse Analysis section above, these explorations are interpretive and therefore make claims that extend beyond Brian and Toby's immediate discursive interactions during activities.

In particular, the analysis considers knowledge practices, broadly construed. Knowledge practices, here, represent knowing as the regularities through which small groups orient to domain practices necessary to participate in Dragon Investigations and feedback conversations. This includes the GenScope software and the language of genetics specifically and of science generally. These knowledge practices develop in and across the aforementioned three small group activities that feature the study participants, Toby and Brian. This study's core goal is

understanding the emergence and transformation of knowledge practices. In contrast to the learning outcome measures above, the framework precludes sharp distinctions between engagement in curricular activities and the learning of introductory genetics. In other words, the enactment of the curricular routines *is* learning and knowing (and vice versa).

Practice is addressed pragmatically by foregrounding epistemic stance during collective participation around the solutions to individual assessment items. Because curricular activities are structured around investigation and assessment items, this analysis develops from *item episodes* as the focal unit of collective practice. Item episodes in and across activities begin to warrant the interpretive claims about the knowledge practices scaffolded by the curriculum and assessment system.

By extension, the learning outcome measures associated with the curriculum also represent learning and knowing. Their distance from the *content* of curriculum and the *context* of its activities, however, render them impoverished accounts of the situated understandings that students, groups, and the class develop. While both learning outcome measures remain meaningful accounts of individual conceptual understanding and proxies for achievement, they are artifacts of situated practice.

Lastly, as a caveat to this analysis, video data come from the natural, and compulsory, setting of a high school. This tension underscores the inherent imposition of schooling and reflects challenges in conducting discourse analytic studies of educational settings. The specific curricular objectives and the general schooling agendas in which these objectives emerge each influence learning environments minute-by-minute and day-by-day. This, in turn, tempers interpretations of a specific learning environment to partial understandings of, at best, probably relevant influences. For these reasons, the scope of this study is narrowly focused on the

participants as they appropriate genetics knowledge practices related to Pedigree Prediction and recursive enactments during feedback conversations.

Procedures

Members of the broader project research team administered the near- and far-transfer learning outcome measures during one regularly scheduled class period several days before and after the implementation. Students who needed additional time or who were absent completed the assessment under the supervision of the classroom teacher.

Between administrations of the learning outcome measures, the classroom teacher, Mr. N, and each of his four life science classes enacted the GenScope curriculum as detailed above. Members of the research team observed all feedback conversations. The author attended the unit 2 feedback conversation only, and considered only video data for the discourse analysis component of the analysis.

The research team also videotaped select activities as each of Mr. N's four classes enacted the curriculum. Two small groups in each class were videotaped during five activities. Because all students engaged in these activities at the same time, the team arranged cameras to capture only the focal group and to exclude all else (see figure 3 below). Video groups sat around circular tables near the corners of the room. Cameras pointed towards corners in order to eliminate background motion and stood slightly above each group in order to capture the color-coded documents being used. The videotaped students also clipped lapel microphones to their clothing in order to reduce background noise. These measures served to foreground the activities of the focal group while, in effect, "deleting" the activities of other groups and the teacher. A qualification in this analysis then is that the collection of video data has, in large measure, deleted the classroom context from the group context.



Figure 3. Sketch image from a videotaped activity

CHAPTER 4

RESULTS

This chapter presents analyses of learning in terms of empiricist, rationalist, and socioculturalist views of transfer and assessment. In the first part, learning is briefly characterized in terms of individual performance on assessment measures that are consistent with rationalist and empiricist views respectively. In the second part, these results and assessment measures are subsumed by analyses consistent with socioculturalist views of learning.

Class-Wide Individual Learning Outcome Measures

Students in Mr. N's four classes gained an average of 19.8 points, or 1.98 standard deviations (SD), on the near-transfer NewWorm performance assessment (NewWorm). These results strongly suggest that participation in the GenScope curriculum impacts understanding of introductory inheritance as it is represented in this assessment. Repeated measures analysis of variance suggest that differences between pre- and post-measures were very unlikely to have occurred by chance, $F(1, 52) = 158.73$, $p < 0.001$. This eclipses gains in two previous implementations in Mr. N's life science classes. In the first year, Mr. N's students gained just 0.65 SD, in the second year they gained 1.52 (Hickey, Kruger, Fredrick, Schafer, & Zuiker, 2003). Further, non-GenScope students who enacted a conventional textbook-based curriculum at Mr. N's school gained just 0.25 SD on the NewWorm. Therefore, results suggest that the present enactments of the curriculum in Mr. N's classes proved to be the relatively more effective at impacting performance on the NewWorm than either Mr. N's previous

implementations or comparison classrooms at the same school. Of course, the NewWorm assessment is biased towards the comparison curriculum by the curricular assessment materials.

Mr. N's students gained an average of 10.6 points, or 1.06 SD, on the far-transfer multiple-choice test. These results also suggest that participation in the GenScope curriculum impacts performance on the multiple-choice test; a repeated measures analysis of variance revealed significant differences between pre- and post-measures, $F(1, 52) = 27.95$, $p < 0.001$. These results reflect trends similar to the NewWorm. They are improvements over other implementations. The present gains scores are three points larger than Mr. N's gains in the previous implementation, though still not significantly different (Hickey, Kruger et al., 2003). This finding is particularly noteworthy because Mr. N's gains nearly double the 0.57 SD gains in the comparison classrooms at Mr. N's school, whose curriculum was relatively biased towards the multiple-choice test. (However, the group x time interaction did not reach conventional criteria for statistical significance ($p < 0.10$, much higher than the usual standard of 0.05) due to the wide variation in gains and the small number of comparison students.)

These overall results fail to reject the first hypothesis of this study: learning about genetic inheritance in the GenScope curriculum transfers broadly to individual performance on conventional assessment practices. As individual performance measures, the NewWorm and multiple-choice test demonstrate learning according to both rationalist and empiricist assessments respectively. As near- and far-transfer measures, they illustrate relative sensitivity to instruction because students obtained larger gains scores on the near-transfer NewWorm than on the far-transfer multiple-choice test. Both characterizations of these measures are useful for demonstrating the impact of GenScope on student learning.

These measures are also necessary to warrant claims about the relative value of the GenScope curriculum for impacting student learning. Nevertheless, this study proceeds under a radically different sociocultural view related to Greeno's theory of situated cognition. The remaining analyses assume that measures of individual performance are peculiar, limited, and thereby incomplete representations (i.e., artifacts) of collective participation in curricular routines. These analyses consider learning primarily in terms of participation and only secondarily in terms of individual performance.

Pedigree Prediction Results of the Learning Outcome Measures

The Pedigree Prediction Dragon Investigation relates to several items on both the NewWorm and the multiple-choice test. Mr. N's classes scored higher on two of three total NewWorm items aligned to Pedigree Prediction. However, class-wide performance on each item shows very small gains. Mr. N's classes also improved on two multiple-choice test items related to Pedigree Prediction, but again these gains were quite small. Therefore, to the extent that these analyses are appropriate for a limited number of assessment items, results for both the rationalist and empiricist assessments reject the second hypothesis of this study: learning about a specific genetic inheritance topic (i.e., making predictions about a pedigree chart) transfers to individual performance on conventional assessment practices related to that topic.

Given the significant overall gains, these lesser gains attest to instructional challenges in impacting learning measures of Pedigree Prediction. As stated earlier, this Dragon Investigation is one of the more difficult topics in the GenScope curriculum and of introductory genetic inheritance curricula in general. It inter-relates three other curricular topics and is therefore relatively complex. It is also an authentic practice of geneticists. These points make it a particularly interesting topic for relating performance and participation. Meanwhile, because the

above analyses suggest that students did not learn about Pedigree Prediction specifically, Pedigree Prediction may be a particularly illustrative topic for profiling the limitations of these measures.

For as much as individual performance measures suggest about the impact of the GenScope curriculum on pedigree predictions, they render incomplete understandings of curricular enactments and student learning. The measures are removed, by degrees, from curricular activities and materials. They under-represent collective participation in curricular routines and therefore necessarily “under-re-present” curricular representations (e.g., GenScope windows).

To make this case, the remainder of this chapter considers one small group’s collective participation relative to the individual performances of its members. As an initial contrast, Brian and Toby’s overall gains, like their classmates, increase on both the NewWorm and multiple-choice test (see table 1 below).

Table 1

Focus Group Students’ and Overall Gains Scores for Learning Outcome Measures

Student	Near-Transfer Gain (SD)	Far-Transfer Gain (SD)
Brian	1.4	0.7
Toby	2.7	2.2
Mr. N’s Classes	2.0	1.1

Their individual performance gains lie above and below Mr. N’s combined classes for each measure however. These differences may relate to differences in individual learning about the domain of introductory genetic inheritance. They may also relate to the degrees to which collective participation during the curriculum transfers to learning outcome measures. Whatever

the case may be, these comparisons begin to relate individual performance and collective participation for the specific set of knowledge practices associated with Pedigree Prediction.

For the purposes of this study, the pre-administration of the measures initiates and the post-administration concludes student engagement with domain practices. On the one hand, the measures are endpoints that measure learning attributable to the enactment of the curriculum that they bound; however, this is only to the degree that items resonate with collective participation in curricular routines. On the other hand, these measures are narrow, isolated representations of knowledge practices that simplify and decontextualize what students learn; however, in this sense, it is only to the degree that items conflict with collective participation in curricular routines. The argument extending from the present conceptual framework suggests that empiricist and rationalist assumptions afford representations that resonate with domain knowledge practices. However, these representations remain too removed from enacted curricular routines to meaningfully represent learning in terms of collective participation (particularly for nascent understandings of challenging topics like Pedigree Prediction).

Analysis of collective participation relative to individual performance necessarily begins by detailing these specific assessment endpoints. These items introduce the practices considered in the interpretive analysis related to the third question of this study and serve as the background against which this analysis will be compared in the fourth and final question of the study. Three items on the NewWorm (out of 24 total) and two items on the multiple-choice (out of 14 total) test relate to the predicting pedigrees using charts. These five items together assess what students learned during the Pedigree Prediction Dragon Investigation and subsequent unit test and final exam feedback conversations.

As part of the near-transfer measure, the NewWorm items related to predicting pedigrees from charts parallel curricular activities. Two items ask students to “read” basic pedigree charts (see items 14 and 15 in figure 4 below). Both assess the concept of dominance relationships, which is featured in unit one and re-appears as part of Pedigree Prediction in unit 3.

PEDIGREE I: DOMINANCE RELATIONSHIPS

Consider two other NewWorm characteristics—Nostrils and Eyes.

- Each characteristic has two phenotypes as shown with the pedigree.
- Females are represented by circles and males are represented by squares.
- Decide what each pedigree says about the dominance relationship between each pair of phenotypes.

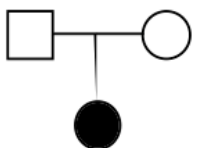
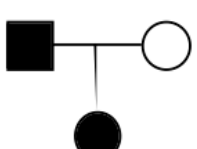
○ Large ● Small □ nostrils ■ nostrils 	14. Having small nostrils is: _____ definitely dominant _____ definitely recessive _____ impossible to tell from what's given
○ Red ● Yellow □ eyes ■ eyes 	15. Having yellow eyes is: _____ definitely dominant _____ definitely recessive _____ impossible to tell from what's given

Figure 4. Items 14 and 15 from the near-transfer NewWorm performance assessment

The pedigree charts in items 14 and 15 show single characteristics and ask students to determine the relationship between parents and one offspring only. In comparison, they resemble the charts featured in the GenScope software and curricular materials (compare with figures 1 and 2 above). Most curricular materials, however, include three generations and many more organisms. They also ask students additional questions that require written responses, like the following

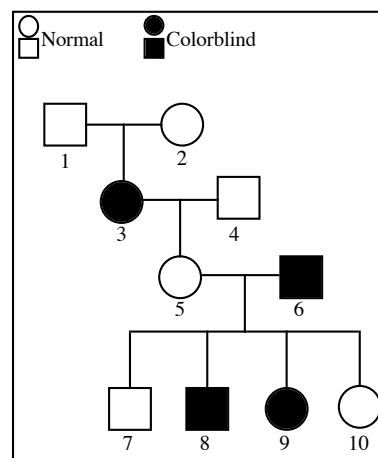
example taken from the third unit test: “Explain what it is about the pedigree that distinguishes between complete and incomplete dominance.” Both of these differences suggest that NewWorm items 14 and 15 appear relatively similar to, but simpler than curricular materials. At the same time, the stems and response alternatives for these items differ from the formulations encountered in the curriculum. They present an additional option in “impossible to tell from what’s given” and all three response alternatives use definitive language (i.e., “definitely” and “impossible”) that is absent in other materials. These differences “distance” the NewWorm items from the curricular materials to some degree. Item 15 is the only Pedigree Prediction item for which students did not perform better; three fewer students answered it correctly on the post-measure than on the pre-measure. The similarities, by the same token, do recognizably present the curriculum in spite of these differences. Similar conclusions can be drawn from a second set of pedigree questions included in the NewWorm.

One NewWorm item with four sub-parts asks students to again read a pedigree chart in order to assess another concept underlying of pedigree predictions called chromosome type. Chromosome type is featured in the second unit and, like dominance relationships, re-appears as part of Pedigree Prediction in unit 3. Item 21 (see figure 5 below) asks students to identify the chromosome type operating across four generations.

PEDIGREE II: COLOR VISION – AUTOSOMAL OR X-LINKED?

Consider another NewWorm characteristic—Color Vision.

- Color Vision has two phenotypes as shown with the pedigree.
- Females are represented by circles and males are represented by squares.
- **Remember:** Males are XX and females are XY.
- Decide if the pedigree is consistent with Color Vision being autosomal or X-linked.



- 21a. Does the Color Vision gene **appear** to be autosomal or X-linked?
 - 21b. Using words and/or diagrams, explain your answer to 21a (use the numbers below each circle or square to refer to particular individuals).
 - 21c. Does this pedigree rule out the type of inheritance you did not choose?
 - 21d. Using words and/or diagrams, explain your answer to 21c (use the numbers below each circle or square to refer to particular individuals).
-

Figure 5. Item 21 from near-transfer NewWorm performance assessment

Some aspects of this item parallel curricular materials. For example, the pedigree chart uses features identical to the Pedigree Prediction Dragon Investigation, unit three test, and final exam, and items 21a and 21b mirror items on the investigation and unit three test as well. Items 21c and 21d, however, present new questions. Both extend from 21b but are never formally addressed in curricular materials. In this way, these latter items advance the problem beyond more familiar ones encountered in the curriculum. Item 21 recognizably presents and extends beyond the curriculum and, together with items 14 and 15, illustrates similarities and differences between the GenScope curriculum and the near-transfer NewWorm performance assessment.

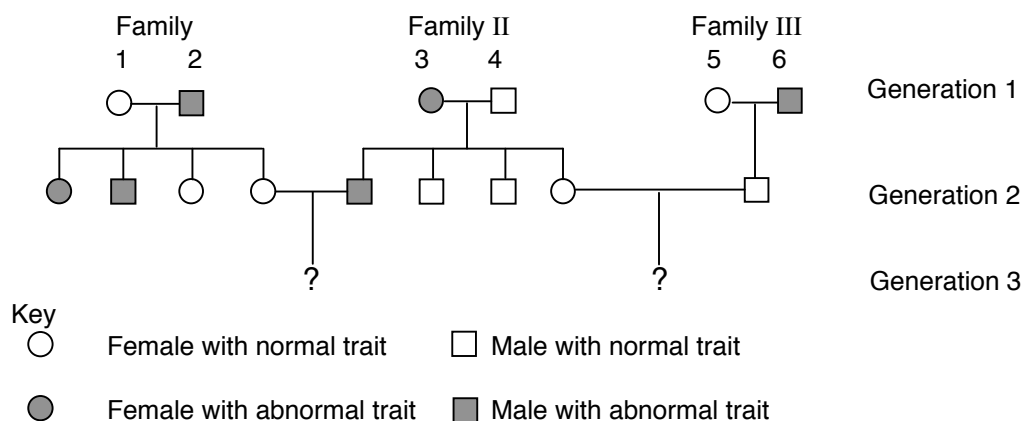
These comparisons between curricular and assessment materials ultimately hinge on student interpretation and performance. Mr. N's students performed better after the curriculum on all items except item 15, however these gains were very small. The focal group, Toby and

Brian, meanwhile did not show gains on any NewWorm items. Brian, however, did answer item 14 correctly before and after the curriculum, and item 21a correctly before the curriculum only. Given that item 21a has only two response alternatives, it is possible that Brian guessed correctly and then incorrectly. Therefore, these performance results demonstrate that Toby and Brian fail to represent any knowledge practices that may have developed during curricular activities. In conclusion, these results reveal only minimal understanding that cannot be attributed to the curriculum and, according the conceptual framework, suggests that neither student will perform better on the far-transfer multiple-choice test.

The much shorter multiple-choice test includes only two items that assess Pedigree Prediction. Both are from the SAT-II item bank. The first item asks students to again “read” a pedigree chart (see figure 6 below). To do so, students must determine the chromosome type for a given trait in ways similar to NewWorm item 21.

Question 13 refers to the diagram below.

The following pedigree traces the appearance of an abnormal recessive trait in several families.



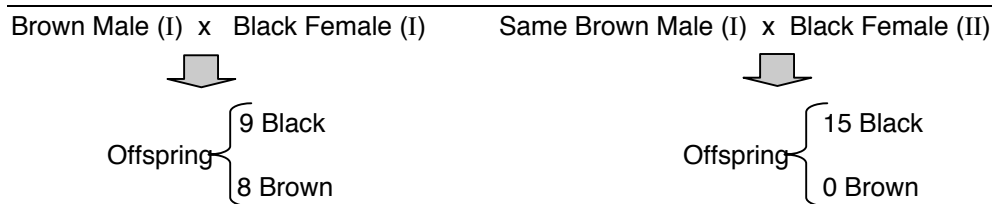
13. Of the following, which provides evidence that the abnormal trait is NOT sex-linked
- (A) Some daughters of individuals 1 and 2 are normal for the trait.
 - (B) Some sons of individuals 3 and 4 are normal for the trait.
 - (C) Half the individuals in generation 1 have the trait.
 - (D) Both sexes have the abnormal condition of the trait.

(E) The trait appears in every generation.

Figure 6. Item 13 from far-transfer multiple choice test

The pedigree chart in multiple-choice item 13 is fairly similar to curricular materials, but includes additional details that mark generations, families, and unknown offspring of a third generation. The question is similar to NewWorm item 21d because both ask why the chart precludes one chromosome type. Item 13 frames a narrower question however. The item stem states the answers to would-be NewWorm questions 21a and 21c, focusing the question on one component of chromosome types. It also further abstracts the problem by omitting organisms (e.g., dragons, worms) and specific characteristics (e.g., eye color, color vision). These variations are not simply differences; relative to the curriculum, they are nearer and farther representations of knowledge practices. Therefore, while item 13 still addresses key aspects of Pedigree Prediction, it does so from a distance less sensitive to the constraints and affordances of knowledge practices emerging in GenScope activities.

A second item makes this point more clearly. Multiple-choice item 14 asks students to read an altogether different kind of pedigree chart (see figure 7 below).



14. One brown male guinea pig was crossed with two black females as illustrated in the diagram above. Which of the following statements is most likely to be true?
- (A) All three parents were heterozygous.
 - (B) All three parents were homozygous.
 - (C) Female (I) is heterozygous.
 - (D) All progeny of female II are homozygous.
 - (E) All black guinea pigs are heterozygous.

Figure 7. Item 14 from far-transfer multiple choice test

The chart diagrams relationships using arrows and brackets together with data rather than shapes, shading, and lines. It assesses the confluence of dominance relationships, allele relationships, and chromosome types, called predicting genotypes, providing sufficient information for students to label the organisms as either homozygous or heterozygous. A smaller but still potentially significant difference is the use of the word progeny in response alternative D. The term “progeny” is not used in the curriculum and therefore likely to be a new and unfamiliar word for many students. Without knowing that progeny is equivalent to “offspring”, which is featured in the GenScope curriculum, Mr. N’s students cannot evaluate one of the five alternatives. For these reasons, item 14 also presents a far-transfer assessment of the curriculum and one further still than item 13 above.

As with the NewWorm, comparisons between the multiple-choice items and the curriculum ultimately hinge on student interpretation and performance. Considering student responses to the above multiple-choice items, Mr. N’s students again improved on items 13 and 14, but these gains are not statistically significant. Student performance on the post-measure parallels the larger population of students who completed these SAT-II biology items. Data provided with these items reveals that only 19% of SAT-II test-takers responded correctly to item 13 compared to 23% of Mr. N’s students. Sixty-six percent of SAT-II students responded correctly to item 14 while only 32% of Mr. N’s students did the same. Meanwhile, Toby and Brian, in particular, both showed gains on item 14, and Brian on item 13 as well. This is a curious outcome inconsistent with the assumptions detailed in the situative framework.

In terms of situated cognition, the constraints and affordances associated with the NewWorm should scaffold knowledge practices more so than the multiple-choice test. Toby and Brian failed to demonstrate competency with NewWorm items relatively similar to the problems they engaged during the curriculum, yet did demonstrate competency with relatively dissimilar multiple-choice items. Table 2 below summarizes these gains for all five items related to Pedigree Prediction.

Table 2

Focus Group's Trajectories for Pedigree Chart Items on Learning Outcome Measures

Student	Measure	<u>Near-Transfer</u>					<u>Far-Transfer</u>		
		14	15	21a	21b	21c	21d	13	14
Brian		*		-				+	+
Toby									+

* = known (right response each time)

- = loss (right to wrong response)

+ = gain (wrong to right response)

These summary results reflect the limitations of mapping student understanding to the theoretical boundaries of transfer delineated in the literature review. Student performance on an assessment cannot exclusively be determined by curricular materials. Each enactment of curricular routines is unique, be it across different small groups and classes, or even the same teacher across classes. As will be shown in the analyses below, the knowledge practices emerging in each group do not necessarily align to the formal domain of genetic inheritance or to representations of it in the learning outcome measures. In short, the curricular materials bound the leaning outcome measures, but not students, classes, and teachers.

This section also considered the alignment of Pedigree Prediction-related assessment items to curricular materials, profiling similarities and differences between materials and items and between separate items. These profiles revealed features of each item that distance them

from the curriculum. They also mentioned three separate aspects of predicting pedigrees called dominance relationships, chromosome types, and allele relationships. These considerations frame the interpretive analysis of Toby and Brian's activities that follows in the next section.

In summary, the results from the learning outcomes measures answer the two hypotheses presented at the outset of the study. First, Mr. N's classes achieved significant learning gains on both general measures of learning gains, the NewWorm performance assessment and the multiple-choice test. Second, Mr. N's students failed to achieve significant learning gains specifically related to Pedigree Prediction on either learning outcome measure. In the cases of Toby and Brian, both students demonstrated overall gains. They also demonstrated specific gains on pedigree-related items while also suggesting a seeming incongruence. Both Toby and Brian performed better on the far-transfer measure than they did on the near-transfer measure. Therefore, at the classroom level, the curriculum impacted student performance in general but not pedigree predictions in particular. At the individual level, Brian and Toby's performance demonstrate gains in general and on far-transfer items related to pedigree predictions.

Discourse Analysis of Collective Participation in Pedigree Prediction Activities

In this section, Toby and Brian's overall performance on both assessments and their specific performance on Pedigree Prediction items are considered in light of their participation across relevant curricular activities. Toby and Brian's discourse during the Pedigree Prediction Dragon Investigation and subsequent feedback conversations demonstrate some consistent knowledge practices that begin to account for the assessed differences detailed above. It is through curricular routines that knowledge practices emerge and later appear in student performance on learning measures.

Accounting for the ways in which students develop knowledge practices during Dragon Investigations and feedback conversations necessarily extends beyond the boundaries of paper-and-pencil assessments. Discursive practices during small group activities trace knowledge practices back to the curricular routines in which they emerged. Therefore, to further make sense of student learning during the GenScope curriculum, the analysis below considers the third and fourth questions in this study, which ask what forms of collective participation students develop and refine and how they relate to conventional assessments of individual performance. Pedigree Prediction practices are considered through the interpretive lens of discourse analysis. In particular, the analysis focuses on discourse that expresses students' ways of knowing about the domain of introductory genetics practices and the social interactions in which these practices occur.

Analyses of Pedigree Prediction Knowledge Practices

GenScope activities encompass complex social interactions framed by explicit curricular intentions. The enactment of these intentions is shaped by curricular scaffolds that arrange a particular set of domain knowledge practices. These scaffolds include the materials, test review steps, and GenScope software detailed above. At the same time, each class uniquely shapes its own enactment. No collective engages curricular scaffolds in exactly the same way. The enactment of a curriculum is therefore a relational endeavor. It is a dynamic interplay between classes and curricula. In this way, each enactment of the curriculum stands as some kind of interactional accomplishment because each enactment pursues meaningful goals, albeit not always aligned with curricular intentions.

The situative framework of this study precludes sharp distinctions between engagement in curricular activities and the learning of introductory genetics. To reiterate an earlier statement,

enacting curricular activities is both learning and knowing (and vice versa). This is because enactments are always the learning of something or the making of some kind of meaning, though not necessarily the intended meanings. The analysis therefore assumes that the learning outcome measures above extend curricular enactments and, as will be addressed through the final question of the study, do not necessarily correspond to Toby and Brian's performances on these measures.

The four analyses below examine discourse around three knowledge practices - dominance relationships, chromosome type, and allele relationships - and their confluence as a fourth knowledge practice, termed predicting genotypes. Stemming from the conceptual framework, each analysis focuses on collective participation, setting aside the concepts and skills featured in pedagogy and assessment. Each section begins with a general profile of collective participation for that specific knowledge practice, which is then supported by student discourse in item episodes.

Analysis 1: Dominance Relationships

Based on analysis across the aforementioned curricular activities, Toby consistently identifies dominance relationships correctly and explains his answers using the given pedigree information for each problem. This suggests that he has enacted domain knowledge practices. Toby formulates legitimate meanings about dominance relationships using pedigree charts and, further, frames those meanings using claim-warrant structures of scientific argumentation. In several item episodes, he also restates his reasoning after reading answer explanations. Brian, on the other hand, generally identifies only some relationships correctly and, although he attempts to explain why, his explanations confound elements of the domain. As shown below, he appropriates irrelevant domain concepts and pedigree information to formulate arguments, suggesting fundamental misunderstandings of domain concepts and their interrelatedness. These

differences in Toby and Brian's formulations mediate their collective participation in dominance relationship problems and individual performance on assessments. This section illustrates these differences using excerpts from group discourse (a transcription key for these excerpts appears in Appendix G).

In the unit three feedback conversation, Toby and Brian recognizably enact the first three test review steps to reconsider an item related to dominance relationships. The item uses a pedigree chart to consider a dragon characteristic called “visual ability” (see item 1 in Appendix B) for which the chart indicates complete dominance. As the excerpt begins, Toby reads the question then asks Brian to state his answer.

- T: Alright (2.0) Alright and one point two says Explain what it is about the pedigree that distinguishes between complete and incomplete dominance. What ju get
- B: I got ((yawns)) wo:: I got um to me I think complete dominance shows for the pedigree complete dominance shows like um x-linked and incomplete shows obstimal ((autosomal)) or you know not dominant because um because its like hold on the majority the majority of the people in there can see instead of bein' blind and half of 'em in mine is are blind.
- T: I:: got I got something different. I got when its completely dominant it's either all the way shaded or all the way plain and if its incomplete dominant it'll be halfway shaded or halfway plain. That's what I got.

In stating his answer, Brian formulates multiple explanations. He first makes two incoherent claims about dominance relationships: (1) complete dominance is x-linked and (2) incomplete dominance is autosomal. He then pauses, saying “hold on,” and proceeds to formulate a third, new, but still incoherent claim: a majority can see but half are blind. This excerpt does not make clear Brian's exact misunderstanding, but it clearly indicates that one exists. Toby's immediate response here is to gloss over Brian's explanation. He simply observes that he “got something different” then proceeds with his own explanation. Toby relates the shading of organisms in

pedigree charts to dominance relationships, which is useful for solving these problems. Together, the formulations in this excerpt illustrate Toby and Brian's disparate understandings.

It remains unclear, however, if the group resolves these differences. Toby's response above and the subsequent reading of the item's answer explanation are possible avenues for developing shared understandings. As such, both represent the possibilities of feedback conversations because each can potentially serve formative functions within the assessment system. Either one can advance collective participation towards increasingly competent domain practices. In this case however, the group does not pursue either one. They simply proceed to the next item without meaningfully accounting for their differences.

The *Judge Your Understanding* rubric is another technology within the assessment system that scaffolds shared understandings. The rubric requires students to explicitly state whether or not they understand unit test items. While discussing the rubric for the item above, Brian states that he, in fact, does understand dominance relationships and Toby does not challenge him. Viewed in terms of curricular objectives, the statement is suspect. Brian does not provide an account of his earlier misunderstanding nor his present claim to understanding. Further, during subsequent final exam item episodes, he again talks about dominance relationships in terms of a majority trait.

This item episode demonstrates some promising possibilities arranged by feedback conversations. It also underscores the need to increase the likelihood that students engage these possibilities. Toby and Brian enact fundamental aspects of scientific argumentation like turn-taking and claim-warrant structures. As a curricular routine however, their collective participation in argumentation does not serve the intended goal of developing shared understandings with each other or with curricular materials. At best, Toby's oblique reference -

that he “got something different” – could be viewed as the initiation of argumentation. More likely however, Toby and Brian’s differences create meaningful bases for argumentation and formative feedback, but the group does not use them as such. The feedback conversation structure, in this instance, fosters collective participation in the recitation of item responses and answer explanations only. The group does not seek consensus and therefore fails to engage in the critical fourth step of the Test Review Steps.

This item episode also clearly illustrates disparities between Toby and Brian’s knowledge practices profiled above. These profiles can be related to each student’s individual performance on NewWorm items 14 and 15. (No multiple-choice items assess dominance relationships using pedigree charts.) Brian’s profile suggests that he does not understand dominance relationships even after the last curricular activity. However, as shown in table 2, Brian’s performance on the NewWorm items suggests otherwise because he responded correctly to item 14 each time. These results are inconsistent with Brian’s demonstrated practice and illustrate a peculiar limitation of measuring individual performance. His performance suggests a degree of learning not demonstrated by his participation in three Pedigree Prediction-related group activities. Toby, on the other hand, recognizes dominance relationships in charts and explains how he recognized them. He does so across all three activities and, further, demonstrates these practices by re-engaging the content as it relates to Brian’s formulations. It is equally inconsistent, then, that Toby failed to improve on NewWorm items 14 and 15 (see table 2). Toby’s formulations across item episodes indicate competence with dominance relationships, however they fail to transfer to similar problems like NewWorm items 14 and 15.

Analysis 2: Allele Relationships

As with dominance relationships, Toby consistently identifies allele relationships correctly. However, he also consistently provides incomplete explanations as to why and may confound practices related to allele relationships with others related to dominance relationships. Brian meanwhile fails to recognize allele relationships correctly and displays fundamental misinterpretations of given pedigree information. Ultimately, neither student demonstrates competent practice alone or together.

To warrant these assertions, the excerpts below draw on the initial investigation and the Unit Three Test feedback conversation. During the investigation, both Toby and Brian simply determine allele relationships by simply tallying the distribution of each trait. Tallies reveal the underlying probabilities and thereby the allele relationships for sufficiently large populations. In this case however, the pedigree chart provides an insufficiently small number of offspring, obscuring the probabilistic operations in the problem. In the following excerpt from the investigation, Toby responds to an investigation question for which he has just looked at a GenScope pedigree chart while counting aloud the number of offspring with each trait.

T: Is having scales dominant, recessive yes, having scales is dominant ((both T and B begin to write on paper))

B: oh yeah ((as if in realization))

T: uh huh because there's more scaled dragons than no scaled dragons

Toby formulates a claim then a warrant, divided by Brian's affirmative "oh yeah." Brian's brief acknowledgement aligns him with Toby's claim, which Toby in turn uses to formulate his warrant. The brevity of this episode is illustrative. Brian's affirmation indexes an indeterminate amount of the group's interactions. It could relate to the whole of Toby's demonstrated practice

of tallying or simply his claim that scales are dominant. Item episodes across all four analyses often remain equally short and obscured by indexicality. Unique to this item however, Brian re-opens the episode after a brief distraction. In the excerpt below, he concedes that the allele relationship remains uncertain.

- B: man I don't know (3.0) I'll try that. I don't know what I'm puttin' (2.0)
 T: yes the majority of offspring ((reading B's written response)) (.) that's just good (.) there are more scaled dragons than non-scaled dragons ((reading from B's paper again))
 B: Is the gene for scales sex-linked?

In response to the uncertainty, Toby verifies that Brian's written answer is consistent with his own, saying "that's just good." Brian then reads aloud the next question without accounting for his uncertainty, thereby ending the item episode. As with dominance relationships, the group enacts aspects of scientific argumentation without developing a shared understanding of domain practices.

Both Toby's misappropriation of tallies and Brian's confusion persist in the unit feedback conversation two days later. In the item episode below, both "read" the pedigree chart and conclude that the blindness allele is dominant (see item 2 in Appendix B).

- B: Ah to me I think the circle part blinds blindness takes over the majority of the offsprings and mine are the offsprings are sighted and the ones that are not circled sighted is dominant and blindness is recessive.
 T: Yeah that's what I got too I think its dominant because there are more blind dragons than sighted.

In this excerpt, Toby warrants his claim by noting a majority, which is consistent with his practice of tallying the expression of each trait. Brian meanwhile misinterprets a superimposed

circle intended as a cue for solving the item. Brian incorporates the circle into his explanation, creating a convoluted formulation.

Brian makes two invalid claims. He argues that allele relationships are different inside and outside the superimposed circle. In making this distinction, he implies that genetic inheritance operates inconsistently. This misinterpretation and his subsequent conclusions suggest that he maintains fundamental misunderstandings. Toby's response is equally troubling. Not only does he fail to recognize Brian's egregious practice, he construes his own altogether different response as similar, if not the same. Immediately following Toby's account, Brian proceeds to read the answer explanation without comment or interruption then the group continues to the next item.

In this episode, Toby and Brian's collective practice is as misleading as in the investigation. Not only do both boys maintain misleading practices about allele relationships, they also foster misleading argumentation practices. They reconcile incommensurable viewpoints by characterizing them as similar without further discussion. This diminishes the formative functions of the feedback conversations to a degree that impedes their participation in domain knowledge practices while also constituting the enactment of the fourth review step.

Interestingly, a few minutes later in the activity Toby seems to critique at least his own understanding. He comments that he cannot explain allele relationships during the *Judge your Understanding* self-assessment. However, evidence from the final exam feedback conversation suggests that his misunderstanding remains despite recognition. Toby again employs the tallying practice in his formulation during the relevant final exam item episode, using a "majority" to warrant his claim. Brian, on the other hand, appropriates the same tallying practice to warrant his claim in the same final exam item episode, contrasting his fallacious claims above. It is ironic

then that, while both students maintained misleading practices, the student who failed to recognize his misunderstanding developed recognizable refinements to his arguments while the student who did recognize his misunderstandings seemingly failed to capitalize on his insight. In this case, it seems probable that Toby's practice provides formative feedback to Brian. Given how little of these practices was spoken, it is likely that feedback remained a largely implicit part of their collective practice, or occurred beyond small group activities.

In summary, Brian refined domain practices related to allele relationships. He fails to explain anything during the investigation, forwards a fundamental misinterpretation in the unit test conversation, but by the exam conversation, provides an explanation that the group, at least, agreed was right. Insofar as Toby's consistent practice is recognized as correct, Brian's appropriation of that practice suggests legitimate appropriation according to the formative assessment model. Meanwhile Toby remains consistent, providing the same explanation in all cases with no observable refinement after two feedback conversations. By the end, they both employ the same tallying practice to determine allele relationships.

None of the assessment items directly assesses allele relationships using pedigree charts. Multiple-choice item 14, however, does assess it indirectly as part of predicting pedigrees. As table 2 indicates, item 14 suggests that Toby and Brian both learned about allele relationships in the GenScope curriculum. The chart provided in the item tallies over thirty offspring, which sufficiently characterizes the probabilistic operations underlying the pedigree. Therefore, each student's gains on this particular item with this particular pedigree chart are consistent with their demonstrated practices in the exam feedback conversation. Toby's consistent formulations and Brian's final one suggest that the practices would transfer to the learning outcome measures and, though indirect, the results for multiple choice item 14 are consistent with that.

Analysis 3: Chromosome Type

Chromosome type is the third basic knowledge practice operating in Pedigree Prediction. The brief profiles that follow are again warranted by item episodes below. Brian identifies but fails to explain how an identified chromosome type derives from information in a pedigree chart. He often appropriates explanations from preceding item episodes or altogether unrelated and seemingly irrelevant other practices. Initially, Toby fails to explain or even identify chromosome types too. He recognizes that he does not understand however, and by the final exam identifies correctly and begins to formulate explanations, albeit misleading ones.

The Pedigree Prediction Dragon Investigation suggests that neither student can explain why a pedigree chart suggests either of two chromosome types: autosomal or x-linked. The excerpt below illustrates the group's initial practice.

B: yeah (.) the gene for scales is sex-linked yeah (3.0)

T: How you know? (2.0)

B: I don't know (2.0) cause because um: because the kids () both the kids um have scales

Brian makes a claim, and fails to formulate a warrant that Toby promptly solicits. Brian cannot explain but restates Toby's explanation from the immediately prior but unrelated item episode. Brian fails to formulate a relevant response, but the group, collectively, engages in effective argumentation. Toby challenges Brian's unwarranted claim then Brian provides one. These beginnings are interrupted however, and the group does not take up this item again. After another chromosome type question appears, Toby identifies the chromosome type but recognizes that he cannot formulate a warrant. The group does not consider chromosome type further during the investigation, but progresses in subsequent feedback conversations.

The Unit Three Test feedback conversation clearly illustrates the group's limited understandings related to chromosome types, and the potential of feedback conversations. The relatively long item episode below displays some of the students' most ill-contrived warrants, but best argumentation.

- T: Alright (1.0) Is the gene for visuality for visual ability autosomal or X-linked? Ah:: I got (.) x-linked.
- B: I got autosomal.
- T: Coz remember in dragons its reversed XX means female and XY mean I mean XY means female and XX means males and its (1.0) more wait hold up
- B: But its sayin that the gene for visual like for sighted is autosomal () but for the circle its autosomal its not x-linked because everybody in there there's only two people out of five that is sighted and three out of five is is blind (1.0) and it says is the gene for visual ability autosomal or x-linked. If it were x-linked then the other two would be blind two.
- T: Oh alright
- B: So its autosomal
- T: Okay
- B: You go ahead (mumbled)
- T: Explain what is it about the circled part of the pedigree that tells you the answer, and explain why the parents and offspring that were not circled do not tell you the answer. Basically what you said (1.0)
- B: Alright=
- T: =cause I messed up on both of them (laughs)
- B: Oh you gotta read three

Brian again appropriates the superimposed circle into his explanation, much like he did for items related to allele relationships above. Toby meanwhile begins to formulate a warrant rooted in the structure of chromosomes but stops, then later concedes that he “messed up” and eventually defers to Brian's incoherent explanations. The group ultimately volleys piecemeal formulations that do not resonate with the underlying practices related to identifying chromosome types.

More interesting however, Toby and Brian's volleys demonstrate active argumentation. The back and forth works to persuasively position warrants for their competing claims. Toby's

formulation embeds a warrant in the appeal “coz remember...”. Brian recasts the argument with “but its sayin...” in order to feature his own warrant. These phrases work to relate and to persuade, building a nascent, shared understanding of problems related to chromosome type. This is the aim of feedback conversations. Unfortunately, it is short lived. Toby quickly agrees with Brian, though taking a seemingly conciliatory stance in the use of “Alright” and “Okay.” This becomes clearer in the last lines of the excerpt when Toby qualifies Brian’s warrant as only “basically” the explanation, then further qualifying it as alright “cause [he] messed up.” This item episode, in sum, illustrates the potentials of feedback conversations for fostering interactions that develop shared understandings and, at the same time, the potential dangers of feedback rooted only in learners’ misguided formulations.

The final exam feedback conversation provides a useful follow-up to the conclusions drawn about chromosome type in the unit three test. The exam conversation occurs several days after the unit test conversation. In between, students review all three unit tests, giving Toby and Brian opportunities to discuss the curriculum in their small group and through teacher-led, whole class discussions. These activities might avert the potential danger that misguided practices will persist without remediation. Further, the probability that misguided practices are indeed averted may relate to the possibilities for students to enact a meaningful wrong, rather than “stage” a meaningless right. Engaging in argumentation around assessment activities may foster understandings about the curriculum and representations of it that enhance the value of activities that follow after it. To this end, the only final exam item episode related to chromosome type illustrates both the potential for learning and potential dangers of not learning. Toby’s formulation in this episode demonstrates sound practice where as Brian’s remains mistaken.

The group's collective participation in items related to chromosome types illustrate disparities. At the same time, the group enacts scientific argumentation to an uncommon degree, albeit with uncommon dissonance with domain practices presented in curricular activities and materials. Both Toby and Brian struggle to make claims and warrants during the initial investigation. These misgivings persist for both students during the unit test feedback conversation. By the final exam feedback conversation however, Toby appears to have developed competent practice while Brian does not. In contrast, the individual performance measures suggest that, if anyone, Brian has improved. Neither student answered NewWorm item 21 correctly, but Brian did gain on multiple-choice item 13, which had proved challenging for both SAT-II test takers and for Mr. N's classes alike. Similar to the relationship between practice and performance in dominance relationships, the representation of Toby and Brian's demonstrated practice with chromosome type appears incoherent according to the study's conceptual framework. Because Toby is able to warrant his claims with sound practice, it follows that it would readily transfer to performance on, at least, items as similar as NewWorm Items 21a and 21b. By the same token, Brian fails to demonstrate across pedigree-related activities any of the practices underlying his successful performance on multiple-choice item 13.

Analysis 4: Predicting genotypes

Coordinating their conclusions from the three practices above, learners lastly predict the genotypes for each individual organism in a given pedigree chart. Labeling each organism is a complex activity. Students must write the appropriate two-letter allele combination, or genotype, next to each organism on a chart. Given the many required responses, each an intricate problem, these items can be rich sources for argumentation. Further, answer explanations provide the correct predictions but do not explain why, leaving claims and warrants entirely in the hands of

each small group. Toby and Brian never discuss their predictions however. Item episodes related to predicting genotypes are short, in part because Brian explains each time that he forgot to do it and also because reading answer explanations typically occupies a significant fraction of each item episode. Without Brian's responses or an answer explanation, Toby simply states his predictions and notes which ones are correct, then the group proceeds to the next item.

The relationship between collective participation and individual performance for predicting genotypes remains largely underdetermined. Toby predicts most genotypes correctly, but without explication of the practices underlying his choices, and Brian never engages the problems. Nevertheless, whatever practices do develop beyond the activities, they serve each student's performance on multiple-choice item 14. Both students respond correctly on the post-measure, suggesting that each successfully coordinated the three constituent practices above and did so with a chart altogether different from those presented in curricular materials.

Summarizing across all four analyses above, differences between Toby and Brian's overall gains illustrate differences in knowing about introductory genetics. Toby's larger gains suggest that he is a more competent geneticist, and these differences resonate with practices associated with dominance relationships, chromosome type, and predicting genotypes for both students. The conclusion is different, however, for the five Pedigree Prediction-related items. NewWorm results suggest that Toby and Brian failed to learn about pedigree predictions while the multiple-choice results suggest that both students did. NewWorm item 14 and multiple-choice item 13 further suggest that Brian may be proficient to some degree more than Toby with specific regard to pedigree predictions. This appears inconsistent with each student's relative, overall performance and with their developing practices across the three activities considered.

The Science and the Lesson

To reiterate the core assumption of this overarching analysis, classroom activities encompass complex social interactions framed by explicit curricular intentions and a range of domain knowledge practices. The curricular activities around the formative assessment model work to shape these knowledge practices through collective participation in feedback conversations. The individuals in each small group in turn work to shape their own unique enactment of the curriculum. Therefore these curricular enactments are relational endeavors that require ongoing negotiations. As such, the interactional accomplishments that comprise any given enactment serve more than just curricular intentions. For each individual, these curricular activities are the confluence of many other, much broader social and cultural practices. The remaining two sections of this analysis attempt to relate these knowledge practices to a possible architecture of the ultimately indeterminate nature of these broader social and cultural practices.

Toby and Brian's interactions across item episodes are not always oriented towards inquiry, argumentation, or the domain of genetics. Rather than engaging in investigations and feedback conversations, they often seek to display compliance with the minimum requirements established for an activity, a classroom, a school. There may be many different reasons for these displays, but this section seeks only to illustrate ways in which they operate within and across item episodes. The next section considers how tensions to identify (or not) and be identified (or not) with certain displays mediate these displays. One common characterization of these interactions is *procedural display* (Bloome, Puro, & Theodorou, 1989). These displays represent what counts as the accomplishment of a lesson or, in this case, an item or test review step. The previous analyses of Pedigree Prediction support other studies (e.g., Jimenez-Alexandre, Rodriguez, & Duschl, 2000) that found argumentation models such as Toulmin's (1958/2003)

insufficient for understanding classroom discourse. The remainder of this section demonstrates why this is so for the present study of Toby and Brian's collective participation. While students sometimes do more than the minimal requirements of the lesson, they more often do less than authentic scientific argumentation. Thus, while the analyses of knowledge practices above provide evidence that Toby and Brian engage the scientific domain of introductory genetics, the excerpts below demonstrate procedural displays of minimal compliance with the lesson.

As they consider the first item in the Pedigree Prediction Dragon Investigation, Toby and Brian cannot explain why a particular trait is or is not a mutation. The excerpt below shows Toby insisting that they consider why this trait might be a mutation.

- B: yeah it's a mutation=
- T: =but why is it a mutation
- B: I don't know ((mumbles))
- T: That is what he just said he said he said [why is your answer
- B: [man you don't "just" man
- T: He said why is your answer your answer.
- B: Just go along with it

Brian attempts to negotiate minimal compliance. The three arrows reference his increasing resistance and ultimately his insistence that they proceed to the next item without reasoning through why the trait is a mutation. Toby, on the other hand, repeatedly affirms that they must understand why, evoking the teacher's intentions with "he." Further, the above interaction occurs before they even read the investigation that asks them to provide a warrant, showing Toby's ongoing desire to display compliance with the lesson. When Brian eventually reads the second part of the question, in which the group is asked to explain, Toby's orientation shifts to the minimal compliance that Brian sought.

- B: How can you tell?
 ⇒ T: It just is. Gotta problem with it?

As the double-line arrow references, Toby shifts from pursuing “why” to mirroring Brian’s opposition, insisting there is no warrant and then issuing a rhetorical challenge to any would-be inquirer (which is clearly not Brian). This exchange illustrates that while Toby’s turns initially work to align the group interactions with activity goals, he soon agrees to realign towards basic procedures. These changes resonate with the notion of procedural display that counts as “doing the lesson” independent of learning goals. Such opposition towards curricular activities reflects efforts to negotiate a less demanding lesson for which they can simply comply. In this sense, Toby and Brian counter the script that is scaffolded by the investigation. Procedural displays illustrate shifts in what collective participation aims to achieve, but it does not account for why these shifts may occur.

Identity in the Science and the Lesson

Jimenez-Aleixandre, Rodriguez, and Duschl (2000) contend that argumentation cannot be understood in terms of models alone, but also procedural display. The analysis of Toby and Brian’s enactment suggests that social identification (Wortham, 2004) also mediates their shifting orientations to the lesson and to the science. Differences in the knowledge practices presented above mediate and shift Toby and Brian’s collaboration in ways that sometimes make visible their efforts to position themselves relative to one another. This section considers the group’s use of epistemic stances for socializing Brian, in particular, to assume the role of a competent geneticist in order to move away from procedural displays and towards authentic inquiry and argumentation.

As a linguistic resource for socializing collective participation, Toby use epistemic stance to construct and highlight Brian's successes during activities related to pedigree predictions. It is often the case that only after negotiating their relative positions during the item episodes that Toby and Brian's collective participation embodies curricular intentions, illustrating the kinds of interactional achievements required to scaffold worthwhile collaborative learning through argumentative discourse. The episodes below demonstrate efforts to ratify such an identity-in-practice for Brian.

A series of evaluations across item episodes in the unit three feedback conversation demonstrate Toby formulating epistemic stances in the service of collective participation. In the first excerpt, both Brian and Toby have stated the same claims but different warrants. Brian's warrant is incoherent and probably incorrect while Toby's matches the answer explanation. Toby does not critique Brian's warrant. In fact, in light of the larger item episode, Toby shows benign neglect for Brian's argument, deferring feedback to the answer explanation. The transcript begins with Toby reading the answer explanation, then foregrounding Brian's claim as "right."

T: Alright From Offspring Phenotypes to Mode of Inheritance. Number one. There are only two phenotypes for visual ability. This suggests complete dominance. So **you** were right - number 1. (Toby points to item 1.1 on Brian's unit test)

B: Yeah

Using "you" and not "we" foregrounds Brian's answer and backgrounds the existing difference. Toby then continues to read the answer explanation, pausing once to point to Brian's pedigree chart while summarizing the answer explanation. This excerpt is the first in a series in which Toby's stance foregrounds Brian's competence. Together with utterances in later item episodes,

Toby's language appears deliberate. In an episode soon after, Toby's stance again evaluates Brian as competent.

- T: ... the gene must be autosomal. Alright so, **you got it. You got it.**
 B: I'm right?
 T: Yeah. **You got it.**
 B: ()
 T: The reason that the uncircled relationships do not help you is because they are autosomal dominant or x-linked dominant inheritance. **You got that.** Alright number 4
 B: Alright

Toby's emphatic stance acknowledges Brian's right answer. He states that Brian "got it" twice then repeats it a third time when Brian's question - "I'm right?" - seeks further affirmation. And again after reading the answer explanation, Toby summarizes that Brian understands. The repetition indexes a second, underlying epistemic stance: Toby ratifies Brian as competent. In effect, he recognizes Brian's participation as valid and meaningful. Toby's repetition thus works to background their differences. He attempts to leverage his competence by simultaneously proposing that Brian "got it" and displaying his alignment with the claim through unnecessary repetition. Returning to the first item exchange, Toby's choice of a singular rather than plural pronoun is consistent with the work he does in the present item episode. The pronoun "you" foregrounds Brian and, just as importantly, backgrounds Toby. Such work serves an important purpose in shaping Brian's participation.

In both episodes, Toby emphasizes elements of the past to preview a possible future, indexing a possible world where Brian is a legitimate student-geneticist of sorts. This pattern illustrates a cultural process called *prolepsis* (Cole, 1996) through which a possible future is projected onto the present. "What is said serves ... to induce presuppositions and trigger

anticipatory comprehension, and what is made known will hence necessarily transcend what is said” (Rommetveit, 1974, p. 88, as cited in Cole, 1996, p. 183). In this way, Toby uses item episodes to construct competence for Brian, which carries over into the next item episode.

In this next item episode, Toby and Brian engage in a relatively complex, protracted argument about dominance relationships. They make oppositional claims, warrant them using data, and importantly, sustain their opposition by agreeing to disagree on the solution. At an impasse, they read the answer explanation. The excerpt begins at the end of the answer explanation where Toby and Brian use an alternative genre to agree to disagree about whether or not the answer is complete or incomplete dominance.

- T: I still say its complete.
 B: Alright - let's take it to the yellow sheet then.
 T: Alright. Fine then. Take it to it then.
 B: Umm: There are only two phenotypes for Fangs. This suggests complete dominance=
 → T: **=com[plete [dominance**
 [[((T hits table with his left hand for emphasis))
 T: **I told you** if it was two ah: choices it would be [complete
 B: **[opp opp opp opp**
 ((B directs gaze and finger at the answer explanation))
 T: wha:t (.) wha:t
 B: I just playin' ((laughs))
 T: ohh ((laughs))

They disagree on who is right and who is wrong and then “take it to the yellow sheet.” This entire item episode signals an appropriation of the activity structure that is neither the typical resistance Brian employs to subvert the lesson nor the compliance they both accede to minimally complete it. Rather, these utterances are an uncommon instance of the use of everyday language to engage in the dialogical process underlying scientific argumentation. As Toby works to legitimate Brian's engagement in domain knowledge practices, Brian increasingly engages them.

They have achieved a sort of positional symmetry in spite of disparities in competent practice by according value to the collective process of argumentation more than either individual's argument. They agree to disagree about the arguments themselves. More importantly however, they agree to privilege their argumentative process rather than their oppositional stances. Such a space and the interactional work to achieve it mediate collective participation in argumentation. As the ongoing item episode demonstrates however, tensions persist in spite of the positional symmetry.

Brian reads the answer explanation, which confirms that Toby's claim is correct. Before continuing on to read the warrants for that claim, Toby interrupts to repeat the key words signaling his victory, as the arrow above references. His inflection and tone invoke the benign positionality he suspended in the item episodes leading to this one. Further, Toby proceeds to explicitly state just how right he is, saying "I told you." Brian interrupts with an overlapping series of "opp's" that counter Toby's shift. It is as if, with these "opp's", Brian holds Toby in check. By alluding to an unread portion of the answer explanation, indexing the prospect of controvertible evidence therein, Brian's "opp opp opp opp" creates doubt, spurring Toby to ask "What? What?" The ruse ends when Brian explains that he was "just playin'," but his innovative epistemic stance succeeds in offsetting Toby's and re-establishes the positional symmetry.

In this final episode, the answer explanation assumes an interesting new role, serving an unintended though useful purpose. As Toby begins to undo the symmetry, Brian employs the answer explanation to leverage a subtle affront. The unread portion of the item's answer explanation plays on Toby's ignorance of the sheet's contents, forcing him to temper his certainty to a more realistic display. The document empowers Brian to challenge Toby with the

believable claim embedded in his “opp’s” and to diffuse the more knowledgeable status that Toby’s enacts.

In summary, by affirming and reaffirming an epistemic stance, Toby’s utterances across item episodes in the unit three feedback conversation demonstrate the mediating influence of more transcendent social and cultural practices. The final item episode represents a different orientation towards argumentation achieved through the fluid, moment-by-moment interactions recurring across multiple item episodes. Utterances like the you-got-it’s operate not only within the differing broader contexts in practice but also around the procedural displays of practice in order to develop a unique interactional accomplishment. The micro-history of practices encompassing pedigree predictions show how Toby and Brian’s enactment shifted from avoiding or minimally complying with (i.e., procedural display) curricular intentions during episodes of the investigation to increasingly engaged though still fragile argumentation.

CHAPTER 5

DISCUSSION

Using three representations of Toby and Brian's learning, the results of this study triangulate the notion of transfer as it relates to educational assessment. Related to the first hypothesis, individual performance on both empiricist and rationalist measures suggest significant learning gains. Overall, they indicated that Mr. N's students achieved statistically significant gains on the near-transfer NewWorm performance assessment and the far-transfer multiple-choice test. Brian and Toby paralleled these results with gains on each measure (see table 1 above). With specific regard to items related to Pedigree Prediction, results show only very small gains for Mr. N's classes on both measures. Toby and Brian failed to show gains on the specific NewWorm items, but did on the multiple-choice items (see table 2 above).

Toby and Brian's collective participation in curricular routines also suggests that they developed some knowledge practices consistent with the domain of introductory inheritance. The four analyses of Pedigree Prediction illustrate Toby's fairly competent domain practices and both students' misunderstandings. In general however, Toby's formulations develop domain practices well beyond Brian's, underscoring his relative competence across nearly all item episodes.

Related to this point, one limitation of the study alluded to in the method section is that video data deleted the classroom context from the group context, providing only partial perspective on Toby and Brian's curricular enactment. Therefore the analyses of curricular practices related to Pedigree Prediction are necessarily incomplete. Gaps exist and therefore, the

extent to which meaningful curricular experiences occur beyond videotaped activities, the profiles are only partial and only partially relate collective practice to individual performance.

The differences in collective participation nevertheless reflect both students' overall gains. The more detailed consideration of individual items suggests a less coherent relationship, however. In two instances, Brian outperformed Toby on learning outcome items in spite of what the group's collective practice suggests about Toby's relative proficiency. The situative conceptual framework employed to cohere the renderings of and relations between these assessment frames represents an innovative perspective for considering the assessment of learning. It demonstrates limitations of the assessment practices that both the NewWorm and multiple-choice test embody. Both serve useful purposes in comparing large populations, but purposes that also necessarily preclude their use as measures of the unique interactional accomplishments and shared understandings that develop (or not) across the enactment of a curriculum. As the four analyses show, assessment items do not align well to the practices developed collectively through participation in curricular activities.

This study is useful because it considers the individual learning outcome measures not only as transfer experiments but also as actual student experiences that differ by degree from curricular experiences. In this way, the analyses responds to Lave's (1980/1997) call to move transfer studies out of the laboratory and into the classroom in both a physical and analytical sense.

This study is also useful because it illustrates the functional limitations of different assessment practices. Acknowledging these limitations and identifying the unique value and service various assessment practices play in education (e.g., formative, summative, accountability) is an important avenue for future research for several reasons. Many research

efforts, for example, increasingly value pointed evidence about the particulars of curricular designs. Reform efforts, on the other hand, demand compelling evidence about the relative impact of competing curriculum. These needs are different and differentially important but each is valuable.

The GenScope curriculum and assessment system demonstrate the different values of each assessment practice. The GenScope curriculum is an effective learning environment. Both the NewWorm and the multiple-choice test warrant this point in terms of learning gains within classes and comparisons between classes (Hickey, Kruger et al., 2003). At the same time, the assessment of collective participation in and through feedback conversations begins to bridge the unique enactments of the curriculum with standardized measures of curricular intentions.

Feedback conversations bridge gaps between the immediacies and particulars of learning environments like GenScope and the generality and abstraction typified by assessment practices like the multiple-choice test, and to a lesser extent the NewWorm. From empiricist and rationalist perspectives, these conversations may appear ineffective and without value because they background the skills and concepts featured in these families of theory respectively. However, the broader project demonstrates the powerful impact feedback conversations can have on learning measures consistent with these perspectives. Further, as the analysis of identity demonstrates, feedback conversations represent a possible site for students to increasingly orient themselves to and to identify themselves with scientific practices.

Future research into feedback conversations must consider the activity structures and, in particular, the social work underlying them. This study suggests that feedback conversations must not only arrange discursive feedback among group members but collective practice as well. The test review steps serve as one example because they structure conversations around

individual assessment responses. These individual stances within collaborative activities may create tensions through the consonance of performance and participation. Specifically, item episodes shape participation around individual arguments rather than collective argumentation, putting the test review steps at odds with the collective orientation of feedback conversations. In only one item episode related to Pedigree Prediction, Toby and Brian accord value to the collective process of argumentation more than either of their arguments. These observations have led to revisions of the test review steps in subsequent projects. Specifically, oppositional language was replaced by more inclusive descriptions (e.g., “sensible solutions” rather than “right or wrong answers”) in order to shift arguments away from polarizing language. This change reflects observations on scientific argumentation by Boulter and Gilbert (1995), who note that “an inclusive – rather than oppositional language has more connection with personal experience” (p. 97).

The review steps also scaffold argumentation within item episodes, but do not sustain these scaffolds across item episodes, activities, or curricula. Feedback conversations, and item episodes in particular, may benefit from scaffolds that support more than the rich formulations of the content, but rich contexts for enacting argumentation. In part, this criticism extends to the GenScope curriculum as well. Neither the curriculum nor the assessment system meaningfully situates the rich content that GenScope generates into a compelling context in which to engage it.

The notion of embodied experience begins to address this distinction between content and context for both assessment (Gee, 2003a) and innovative curriculum designs (Gee, 2003b). An embodied experience loosely describes interactions that are “value-laden” and “perspective-taking” and not simply information or facts. In this sense, Dragon Investigations and feedback conversations arrange opportunities to do genetic science and allow occasions to do the lesson

(i.e., procedural display), but not to *be* the geneticist. Research into feedback conversations that scaffold embodied experiences may support what Greeno, Collins, and Resnick (1996) call “positive epistemic identities.” Designing learning environments that support the development of competent learners and knowers may also scaffold and sustain participation in these learning activities and communities. Research already underway is addressing this point by modifying unit test formats and answer explanations in order to re-structure feedback conversations as more embodied experiences (Zuiker, Hickey, Kwon, & Barab, 2005).

Future research should also extend beyond feedback conversations to the GenScope assessment framework. Pedigree Prediction illustrates the possibilities for a kind of embedded assessment practice. It has double relevancy as both a new practice in the GenScope curriculum and as a formative assessment practice of dominance relationships, allele relationships, and chromosome types. The assessment of these knowledge practices situates them in broader practices of predicting genotypes, thereby increasing participation in the science of genetic inheritance (Lave & Wenger, 1991). This idea describes the notion of *liminal assessment*. Liminal assessments, in concept, are an assessment practice seamlessly embedded into curricular experiences. The format and questions resonate with and leverage the content and context of a curriculum in order to generate compelling and pointed evidence of collective participation. In this way, liminal assessments not only relate to curricular intentions, they are the enactment of those intentions, making visible learner understandings while further engaging them in the curriculum. Liminal assessment integrates practices towards increasing participation in a domain, embedding questions along a trajectory of increasing participation. In this sense, a practice can be liminal to the extent that it serves as the background against which a new practice has meaning.

Lobato (2003) has developed what she terms an actor-oriented perspective on transfer that is similar. In her view, the analytical focus revolves around the unique process of knowing in designated problems rather than expected outcomes (e.g., this study's learning outcome measures). The key difference in a liminal problem is a consideration of the embeddedness of the unique processes within the contexts of increasing participation in communities of practice. Such embeddedness promises to foster collective participation with content and in contexts while also making visible the constituent knowledge practices of a liminal assessment. In this way, liminal assessment begins to develop feedback conversations through collective participation in curricular experiences.

REFERENCES

- Adger, C. T. (2003). Discourse in educational settings. In D. Schiffrin, D. Tannen & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 503-517). Malden, MA: Blackwell.
- Assessment Reform Group. (1999). *Assessment for learning; Beyond the black box*. Cambridge: University of Cambridge School of Education.
- Bakhtin, M. (1998). Discourse in the Novel. In J. Rivkin & M. Ryan (Eds.), *Literary Theory: An Anthology* (pp. 32-44).
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy of far transfer. *Psychological Bulletin*, 128(4), 612-637.
- Bateson, G. (1972). *Steps to an ecology of mind*. San Francisco: Chandler Publishing Company.
- Bateson, G. (1979). *Mind and nature*. New York: E. P. Dutton.
- Beach, K. (1999). Consequential Transitions: A Sociocultural Expedition Beyond Transfer in Education. *Review of Research in Education*, 24, 101-139.
- Bell, B., & Cowie, B. (2001a). The characteristics of formative assessment in science education. *Science Education*, 85, 536-553.
- Bell, B., & Cowie, B. (2001b). *Formative assessment and science education*. Dordrecht, The Netherlands: Kluwer.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Bloome, D., Puro, P., & Theodorou, E. (1989). Procedural display and classroom lessons. *Curriculum Inquiry*, 19(3), 265-292.
- Boulter, C. J., & Gilbert, J. K. (1995). Argument and science education. In P. S. M. Costello & S. Mitchell (Eds.), *Competing and consensual voices: The theory and practice of argumentation*. Clevesdon, UK: Multilingual Matters.
- Bransford, J. B., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61-100.

- Case, R. (1996). Changing views of knowledge and their impact on educational research and practice. In D. R. Olson & N. Torrance (Eds.), *The Handbook of Education and human Development: new Models of Learning, Teaching, and Schooling* (pp. 75-99). Cambridge, MA: Blackwell.
- Chafe, W. L., & Nichols, J. (Eds.). (1986). *Evidentiality: The linguistic coding of epistemology*. Norwood, N.J.: Ablex Pub. Corp.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: The Belknap Press of the Harvard University Press.
- Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education*, 6(1), 101-116.
- Delandshere, G. (2002). Assessment as inquiry. *Teachers College Record*, 104(7), 1461-1484.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287-312.
- Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment*, 4(1), 37-73.
- Flyvbjerg, B. (2001). *Making social science matter*. New York: Cambridge University.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Garfinkel, H., & Rawls, A. W. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Lanham, Md.: Rowman & Littlefield Publishers.
- Gee, J. P. (1999). *An introduction to discourse analysis*. New York: Routledge.
- Gee, J. P. (2003a). Opportunity to learn: a language-based perspective on assessment. *Assessment in Education*, 10(1), 27-46.
- Gee, J. P. (2003b). *What video games have to teach us about learning and literacy* (1st ed.). New York: Palgrave Macmillan.
- Gipps, C. (1999). Sociocultural aspects of assessment. *Review of Educational Research*, 24, 355-392.
- Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17.
- Greeno, J. G., Collins, A., & Resnick, L. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.), *The handbook of educational psychology* (pp. 15-46). New York: MacMillan.

- Greeno, J. G., Smith, D. R., & Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 99-167). Norwood, NJ: Ablex.
- Greeno, J. G., & the Middle School Mathematics through Application Project Group. (1997). Theories and practices of thinking and learning to think. *American Journal of Education*, 106(November), 85-126.
- Greeno, J. G., & the Middle School Mathematics through Application Project Group. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5-26.
- Gumperz, J. J. (Ed.). (1982). *Language and social identity*. Cambridge: Cambridge University Press.
- Hall, R. P., Knudsen, J., & Greeno, J. G. (1996). A case study of systemic aspects of assessment technologies. *Educational Assessment*, 3(4), 315-361.
- Halliday, M. A. K. (1978). *Language as a social semiotic*. London: Edward Arnold.
- Harré, R., & Gillett, G. (1994). *The discursive mind*. Thousand Oaks, CA: Sage Publications.
- Heller, M. (2003). Discourse and interaction. In D. Schiffrin, D. Tannen & H. E. Hamilton (Eds.), *The handbook of discourse analysis*. Malden, MA: Blackwell.
- Hickey, D. T. (2003). Engaged participation versus marginal non-participation: A stridently sociocultural approach to achievement motivation. *The Elementary School Journal*, 103(4), 401-429.
- Hickey, D. T., Kindfield, A. C. H., Horwitz, P., & Christie, M. A. (2003). Integrating curriculum, instruction, assessment, and evaluation in a technology-supported genetics environment. *American Educational Research Journal*, 40(2), 495-538.
- Hickey, D. T., Kruger, A. C., Fredrick, L. D., Schafer, N. J., & Zuiker, S. J. (2003, April). *Design experimentation with multiple perspectives: The GenScope Assessment Project*. Paper presented at the Annual Conference of the American Educational Research Association, Chicago, IL.
- Hickey, D. T., Wolfe, E. W., & Kindfield, A. C. H. (2000). Assessing learning in a technology-supported genetics environment: Evidential and consequential validity issues. *Educational Assessment*, 6(3), 155-196.
- Hickey, D. T., & Zuiker, S. J. (2003). A new perspective for evaluation of innovative science environments. *Science Education*, 87(3), 539-563.

- Holland, D., Lachicotte Jr., W., Skinner, D., & Cain, C. (1998). *Identity and agency in cultural worlds*. Cambridge: Harvard University Press.
- Horwitz, P., & Christie, M. A. (2000). Computer-based manipulatives for teaching scientific reasoning: An example. In M. J. Jacobson & R. B. Kosma (Eds.), *Learning the sciences of the twenty-first century: Theory, research, and the design of advanced technology learning environments*. Hillsdale, NJ: Lawrence Erlbaum.
- Jimenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, 84, 757-792.
- Kelly, G. J., & Duschl, R. A. (2002, April). *Toward a research agenda for epistemological studies in science education*. Paper presented at the National Association for Research in Science Teaching, New Orleans, LA.
- Kindfield, A. C. H., Hickey, D. T., & Yessis, L. M. (1999, March). *Assessing Student Understanding of Genetics: The NewWorm Assessment*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Boston, MA.
- Latour, B. (1987). *Science in action : How to follow scientists and engineers through society*. Cambridge, Mass.: Harvard University Press.
- Latour, B., & Woolgar, S. (1986). *Laboratory life : the construction of scientific facts*. Princeton, N.J.: Princeton University Press.
- Lave, J. (1980/1997). What's special about experiments as contexts for thinking. In M. Cole, Y. Engerstrom & O. Vasquez (Eds.), *Mind, culture, and activity: Seminal papers from the laboratory of comparative human cognition*. Cambridge, UK: Cambridge University.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- Lemke, J. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing Corporation.
- Lemke, J. (1995). *Textual politics: Discourse and social dynamics*. Bristol, PA: Taylor & Francis.
- Lemke, J. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity*, 7(4), 273-290.
- Lightfoot, C., & Cox, B. D. (1997). Locating competence: The sociogenesis of mind and the problem of internalization. In B. D. Cox & C. Lightfoot (Eds.), *Sociogenetic perspectives on internalization* (pp. viii, 265). Mahwah, NJ: Lawrence Erlbaum Associates.

- Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher*, 32(1), 17-20.
- Mayer, R. E., & Whittrock, M. C. (1996). Problem-solving transfer. In D. Berliner (Ed.), *Handbook of Educational Psychology* (pp. 47-62).
- National Research Council. (1999). *The assessment of science meets the science of assessment: Summary of a workshop*. Washington, D.C.: National Academy Press.
- National Research Council. (2001a). *Classroom assessment and the national science education standards*. Washington, D.C.: National Academy Press.
- National Research Council. (2001b). *Knowing what students know : The science and design of educational assessment*. Washington, DC: National Academy Press.
- National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- National Research Council. (2003). *Assessment in Support of Instruction and Learning: Bridging the Gap Between Large-Scale and Classroom Assessment*. Washinton, D.C.
- Ochs, E. (1996). Linguistic resources for socializing humanity. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 407-437). Cambridge: Cambridge University Press.
- Ochs, E., Gonzales, P., & Jacoby, S. (1996). "When I come down I'm in the domain state": grammar and graphic representation in the interpretive activity of physicists. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 328-369). Cambridge: Cambridge University.
- Packer, M. (2001). The Problem of Transfer, and the Sociocultural Critique of Schooling. *The Journal of the Learning Sciences*, 10(4), 493-514.
- Packer, M., & Goicoechea, J. (2000). Sociocultural and constructivist theories of learning: Ontology, not just epistemology. *Educational Psychologist*, 35(4), 227-241.
- Payne, D. A. (1997). *Applied educational assessment*. New York: Wadsworth.
- Roth, W.-M. (1996). The Co-Evolution of Situated Language and Physics Knowing. *Journal of Science Education and Technology*, 5(3), 171-191.
- Roth, W.-M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research*, 71(3), 365-392.
- Roth, W.-M. (2003). Gesture-speech phenomena, learning, and development. *Educational Psychologist*, 38(4), 249-264.

- Sacks, H. (1992a). *Lectures on conversation* (Vol. I). Oxford: Basil Blackwell.
- Sacks, H. (1992b). *Lectures on conversation* (Vol. II). Oxford: Basil Blackwell.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sadler, R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5, 77-84.
- Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right ... if they argue together! *Cognition and Instruction*, 18(4), 461-494.
- Stewart, J., & Hafner, R. (1994). Research on problem solving: Genetics. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 284-300). New York: Macmillan.
- Thorndike, E. L. (1906). *The principles of teaching, based on psychology*. New York,: A.G. Seiler.
- Torrance, H. (1993). Formative assessment: Some theoretical problems and empirical questions. *Cambridge Journal of Education*, 23, 333-343.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham, England: Open University Press.
- Toulmin, S. E. (1958/2003). *The uses of argument* (Updated ed.). New York: Cambridge University Press.
- Turnstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, 22, 389-404.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge: Harvard University Press.
- Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge: Harvard University Press.
- Wilson, M. (Ed.). (2004). *Towards coherence between classroom assessment and accountability*. Chicago: The University of Chicago Press.
- Wortham, S. (2004). The interdependence of social identification and learning. *American Educational Research Journal*, 41(3), 715-750.
- Zimmerman, B. J. (1993). Commentary. *Human Development*, 36, 82-86.

Zuiker, S. J., Hickey, D. T., Kwon, E. J., & Barab, S. A. (2005, April). Framing Quest Atlantis: A multi-level/multi-type assessment framework for learning and assessment. Paper presented at the bi-annual Computer Assisted Learning conference, Bristol, England.

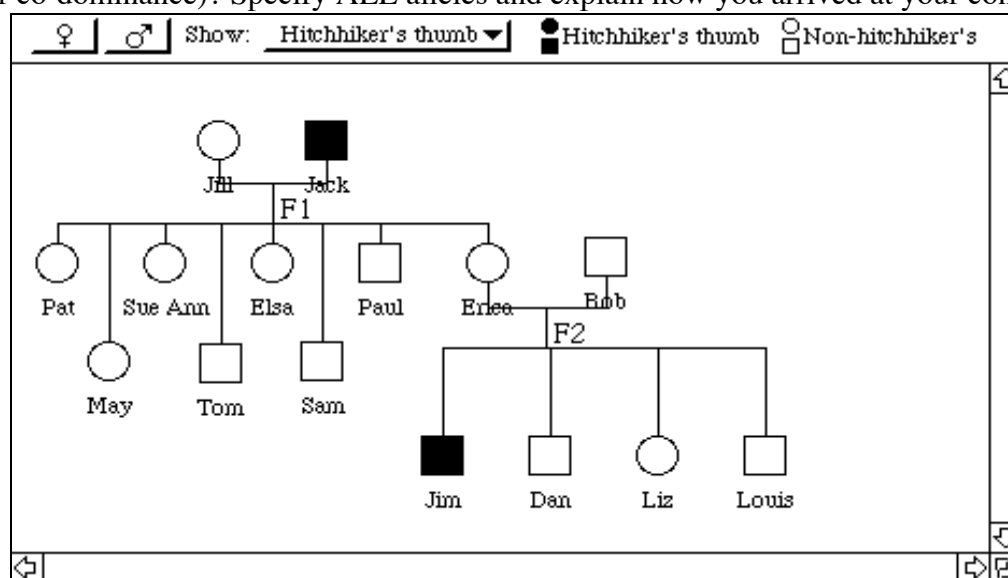
APPENDICES

APPENDIX A

PEDIGREE PREDICTION DRAGON INVESTIGATION

Dragon Investigation: Pedigree Predictions I

1. A male and female dragon, each with spots, produce 100 offspring with the following distribution of phenotypes:
 - 77 with spots (51 males and 26 females)
 - 23 with no spots (0 males and 23 females)
- (a) How might a pedigree for this dragon pair look? (Hint: You do not have to draw 100 circles and squares to answer this question, use percentages.)
- (b) What is the mode of inheritance for spots (autosomal or X-linked; complete, incomplete, or co-dominance)? Specify ALL alleles and explain how you arrived at your conclusion.



This pedigree shows the human trait hitchhiker's thumb for Jack, Jill, their children, and their grandchildren. Jack and his grandson Jim have hitchhiker's thumb. Jill does not nor do any other children or grandchildren.

2. Which of the following is Jim's genotype for hitchhiker's thumb? **TT** **Tt** **tt** **T-** **t-**
3. What must Erica's and Bob's genotypes be? **TT** **Tt** **tt** **T-** **t-**
4. What is Jack's genotype for hitchhiker's thumb? **TT** **Tt** **tt** **T-** **t-**
5. What is Jill's likely genotype for hitchhiker's thumb? **TT** **Tt** **tt** **T-** **t-**
6. If Jack and Jill have 12 more children, how many do you think will have hitchhiker's thumb?

None	3 or 4	About half	All of them
------	--------	------------	-------------
7. If Erica and Bob have 12 more children, how many do you think will have hitchhiker's thumb?

None	3 or 4	About half	All of them
------	--------	------------	-------------

8. Which of the following terms describe the mode of inheritance of Hitchhiker's Thumb?
(circle ALL that apply)

Complete dominance	Autosomal	Dominant trait
Incomplete dominance	X-linked	Recessive trait
Co-dominance		
9. Draw a circle around the individuals in the Jack & Jill pedigree that indicate what type of inheritance this is. Explain your reasoning
10. OK—now try it on GenScope™. Open the file *HHThumb.gs* and check out the chromosomes. Also, give Jack and Jill and Erica and Bob some more kids and see if you made the right predictions. Record your findings below.

Dragon Investigation: Pedigree Predictions II

Independent Practice: This activity will help you to understand a new trait through patterns of inheritance and effect-to-cause reasoning. Open the *ScaledDragon.gs* file and you will see a dragon that you may not have ever seen before. It has scales all over its body! What's going on here? Here are some questions for you to work on as you think about this problem.

1. Is this new trait the result of a mutation or did the scaled dragon inherit scales from its parents? Explain your answer.
2. Is having scales dominant (like horns), recessive (like wings), incompletely dominant (like legs), or something else? How can you tell?
3. Is the gene for scales sex-linked (check one choice only):
 - ☐ Sex-linked (in the X or the Y chromosome; if so, which one)?
 - ☐ Autosomal (in Chromosome 1 or Chromosome 2)?
 How can you tell?

*Open the file *PlatedDragon.gs* and you will see another different kind of dragon. This one has plates on its neck. Now try to figure out the mode of inheritance for this trait.*

4. Is this new trait the result of a mutation or did the plated dragon inherit plates from its parents? Explain your answer.
5. Is having plates dominant (like horns), recessive (like wings), incompletely dominant (like legs), or something else? How can you tell?
6. The gene for plates is (check one choice only):
 - ☐ Sex-linked (in the X or the Y chromosome; if so, which one)?
 - ☐ Autosomal (in Chromosome 1 or Chromosome 2)?
 How can you tell?

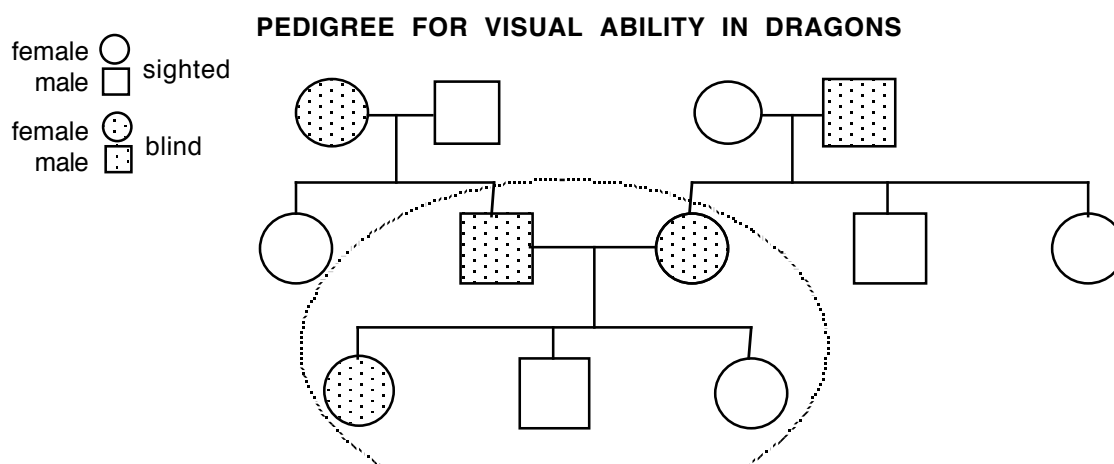
APPENDIX B

UNIT THREE TEST MATERIALS RELATED TO PEDIGREE PREDICTION

Section 3A: Assessment

From Offspring Phenotypes to Mode of Inheritance I

In dragons, a single gene with two possible alleles determines visual ability. Use what you know about pedigrees to help you figure out these other things about visual ability.



- 1.1 Do the alleles for visual ability show **complete dominance** or **incomplete dominance**? _____
- 1.2 Explain what it is about the pedigree that distinguishes between complete and incomplete dominance. _____
- 2.1 Is the allele for blindness **dominant**, **recessive**, or **incompletely dominant** to the sighted allele? _____
(Hint: use the circled part of the pedigree.)
- 2.2 Explain what it is about the circled part of the pedigree that tells you the answer, and explain why the parents and offspring that were not circled do not tell you the answer. _____
- 3.1 Is the gene for visual ability **autosomal** or **X-linked**? _____
(Hint: use the circled part of the pedigree and remember that male dragons are XX and females are XY.)
- 3.2 Explain what it is about the circled part of the pedigree that tells you the answer, and explain why the parents and offspring that were not circled do not tell you the answer. _____
- 4 Write the possible genotypes **on every individual** in the pedigree above.
(Hint: at least one individual has more than one possible genotype.)

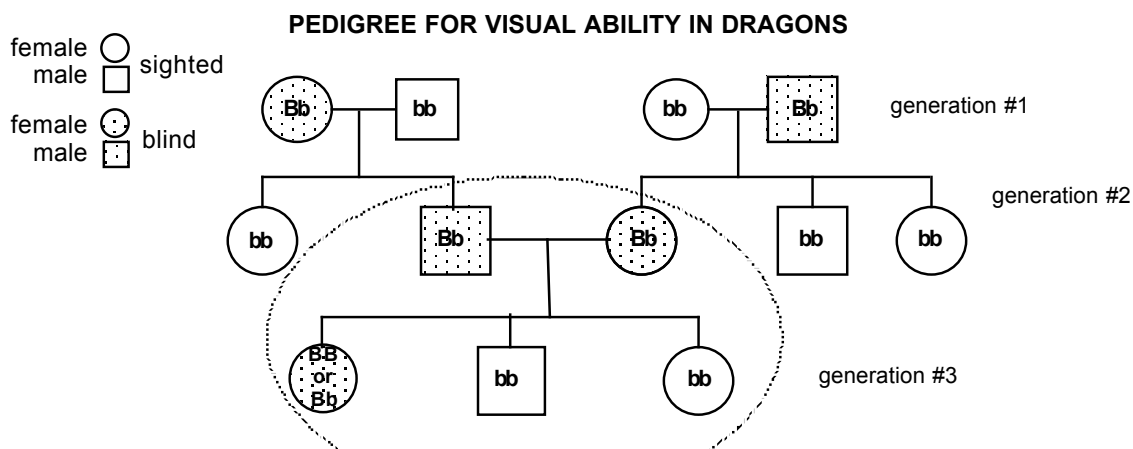
Section 3A: Answer Explanations

From Offspring Phenotypes to Mode of Inheritance I

- There are only two phenotypes for visual ability. This suggests **complete dominance**. If a third intermediate phenotype (like partial blindness) would suggest incomplete dominance. The two phenotypes are represented in the pedigree by completely shaded and completely unshaded circles and squares. If a third phenotype were present, it would have been represented by half-shaded circles and squares. NOTE: The mere absence of any half-shaded circles or squares doesn't **rule out** incomplete dominance but specific clues in the overall pedigree do. Can you find one clue that does?
- The circled part of the pedigree shows two blind parents who have sighted offspring. If the blindness allele was recessive, both parents would be homozygous recessive, and therefore all of their offspring would also be homozygous recessive and blind. Since some offspring are sighted, the blindness allele must be dominant and each parent must be heterozygous. The reason that the uncircled relationships don't help you is because they both involve a blind parent and a sighted parent.
- If the gene for visual ability was X-linked, then only the female offspring of blind parents (in the circled part of the pedigree) would be sighted. Since one of the sighted offspring is male, the gene must be autosomal. The reason that the uncircled relationships don't help you is because they are consistent with either autosomal dominant or X-linked dominant inheritance.

If any of the above are unclear, try drawing Punnett squares assuming different modes of inheritance (autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive) to make sense of them.

4.



Judge Your Understanding of 3A: From Offspring Phenotypes to Mode of Inheritance I

These questions involved **Effect-to-Cause** problems in a **Between-Generation** setting. This means that you had to look at an **effect** (the patterns of inheritance in the pedigree) and figure out what might have **caused** it (the mode of inheritance of the characteristic).

FIRST...you have reviewed the answer explanations for four kinds of problems. Judge whether or not you understand each type well enough to solve similar problems for other organisms.

Type	Question #	Description	Check if you understand
Recognize type of dominance relationship	1.1 & 1.2	Use a pedigree to determine type of dominance relationship.	_____
Recognize other aspects of mode of inheritance	2.1 & 3.1	Use a pedigree to determine specific allele relationships and chromosome type.	_____
Explain other aspects of mode of inheritance	2.2 & 3.2	Explain how a pedigree is used to determine specific allele relationships and chromosome type.	_____
Label pedigree	4	Use understanding of modes of inheritance to determine genotypes of individuals in a pedigree.	_____

THEN ... judge your understanding of pedigree and mode of inheritance. Put a check mark in the box that is your best estimate of our level of understanding.

If your understanding of pedigrees and modes of inheritance is ...	Check One
EXEMPLARY , you can probably solve all four kinds of problems:	_____
ACCOMPLISHED , you can probably solve three of the four kinds of problems:	_____
DEVELOPING , you can probably solve two of the four kinds of problems:	_____
BEGINNING , you can probably solve only one of the four kinds of problems:	_____

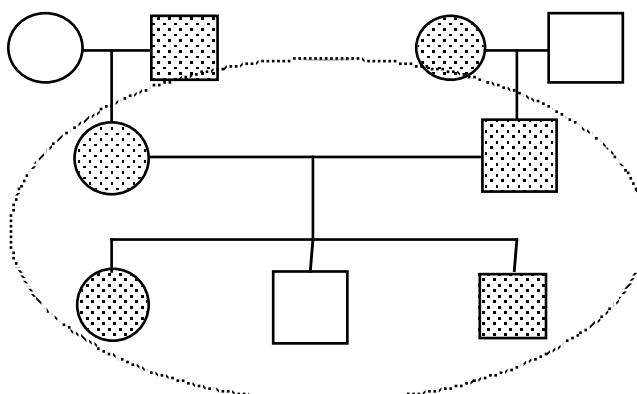
APPENDIX C

FINAL EXAM MATERIALS RELATED TO PEDIGREE PREDICTION

From Offspring Phenotypes to Mode of Inheritance I: Pedigrees

In dragons, a single gene with two possible alleles determines **horn color**. Use what you know about pedigrees to help you figure out more about the mode of inheritance of horn color. Note: all dragons in the pedigree have horns.

female  black
 male  horns
 female  white
 male  horns

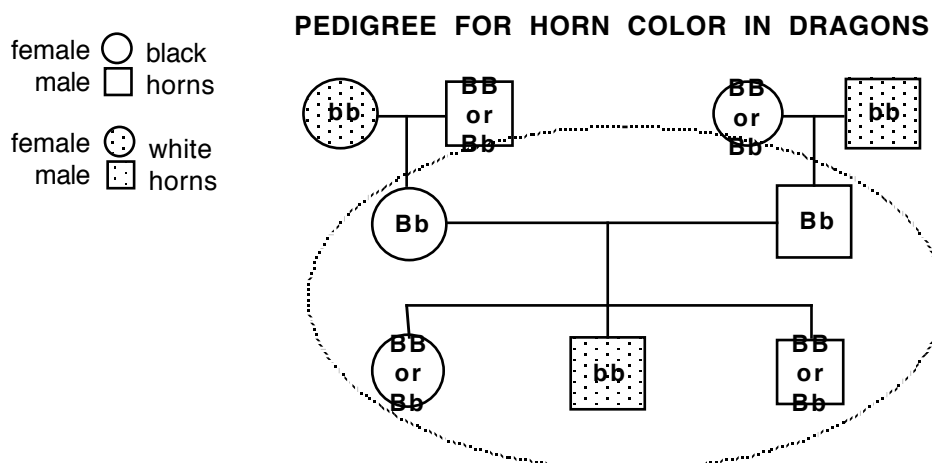
PEDIGREE FOR HORN COLOR IN DRAGONS

- 15a.** Do the alleles for horn color show **complete dominance** or **incomplete dominance**? _____
- 15b.** The allele for white horns is **recessive** to the allele for black horns. Explain what it is about the circled part of the pedigree that tells you that the white allele is recessive.
- 15c.** The gene for Horn Color is **autosomal**. Explain what it is about the circled part of the pedigree that tells you the answer.
- 15d** Using **B** for the black allele and **b** for the white allele, write ALL possible genotypes **on every individual in the pedigree above**. (Hint: at least one individual has more than one possible genotype.)

From Offspring Phenotypes to Mode of Inheritance I: Pedigrees

Answer Explanations (embedded)

In dragons, a single gene with two possible alleles determines **horn color**. Use what you know about pedigrees to help you figure out more about the mode of inheritance of horn color.. Note: all dragons in the pedigree have horns.



- 15a. Do the alleles for horn color show **complete dominance** or **incomplete dominance**? **complete**
- 15b. The allele for white horns is **recessive** to the allele for black horns. Explain what it is about the circled part of the pedigree that tells you that the white allele is recessive.
- The circled part of the pedigree shows two parents with black horns who have a son with white horns. If the white allele was dominant, at least one of the parents would have to have that allele, and therefore would have white horns. Since both parents have black horns, the white allele must be recessive and each parent must be heterozygous (if autosomal) or heterozygous/hemizygous (if X-linked).**
- 15c. The gene for Horn Color is **autosomal**. Explain what it is about the circled part of the pedigree that tells you the answer.
- If the gene for Horn Color was X-linked, then only the female offspring of parents with black horns in the circled part of the pedigree could have white horns. Since the offspring with white horns is male, the gene must be autosomal.**
- 15d. Using **B** for the black allele and **b** for the white allele, write ALL possible genotypes **on every individual** in the pedigree above. (Hint: at least one individual has more than one possible genotype.)

APPENDIX D

TEST REVIEW STEPS

The more you review your unit tests, the better you will do on the final and the graduation test. Spend the entire period reviewing with your group. Use data and knowledge of genetics to support scientific claims.

FOR EACH ITEM:

1. Each student must state and defend a solution. If you don't know—guess!
2. Work together to figure out the best solution. Why is the correct answer correct and why are the wrong answers wrong? Each student should agree on the solution or “agree to disagree” before the next step.
3. Read the answer explanation (the yellow sheet) together. Compare that solution with your own solution(s). Each student should state whether their solution was the same or different to that solution.
4. Make sure that every student understands why the correct answer is correct and why wrong answers are wrong before going to the next item.

AFTER EACH SECTION OF THE TEST:

1. Review the “Judge Your Understanding Rubric” for each type of problem.
2. Each student should state which kinds of problems they think they could solve on another test (like the final). Check the boxes accordingly.
3. According the rubric, each student must state whether their understanding is beginning, developing, accomplished, or exemplary.

APPENDIX E

NEWWORM PERFORMANCE ASSESSMENT

The NewWorm®

Original image copyright 1997, William Wadsworth.
Used with permission.



The left box shows what we know about NewWorms' genes. The right box shows the genetic makeup of two NewWorms. Use this information to solve the problems below.

NewWorm Genetics	One NewWorm Genotype
Body: Flat: BB or Bb Round: bb Mouth: Oval: ?? Slit: ?? Head: Broad: ?? Medium: ?? Narrow: ?? Rings: No Rings: RR or Rr Rings: rr Color: Green: CC Brown: Cc Black: cc Tail (Male): Pointed: TT or Tt Blunt: tt Tail (Female): Pointed: T- Blunt: t- Note: The Tail gene is in the X chromosome. The — indicates that the gene is not present in the Y-chromosome. Sex: Males: XX Females: XY	NewWorm1

GENOTYPE-PHENOTYPE MAPPING

Determine phenotypes (traits) from NewWorm1's genotypes:

NewWorm1

- | | |
|---------------------|----------|
| Does it have rings? | 1. _____ |
| What color? | 2. _____ |
| What kind of tail? | 3. _____ |
| Male or female? | 4. _____ |

If the allele for **broad head (H)** is incompletely dominant to the allele for **narrow head (h)** and **medium head** is in between broad and narrow:

- | | |
|--------------------|----------|
| What kind of head? | 5. _____ |
|--------------------|----------|

NewWorm Genetics	Two NewWorm Phenotypes NewWorm2
Body: Flat: BB or Bb Round: bb	round body
Mouth: Oval: ?? Slit: ??	oval mouth
Head: Broad: ?? Medium: ?? Narrow: ??	medium head
Rings: No Rings: RR or Rr Rings: rr	no rings
Color: Green: CC Brown: Cc Black: cc	green
Tail (Male): Pointed: TT or Tt Blunt: tt	pointed
Tail (Female): Pointed: T– Blunt: t–	
Note: The Tail gene is in the X chromosome. The — indicates that the gene is not present in the Y-chromosome.	
Sex: Males: XX Females: XY	female

PHENOTYPE-GENOTYPE MAPPING

For each characteristic, circle ALL of NewWorm2's possible genotypes.

Characteristic	NewWorm2				
6. Body	BB	Bb	bb	B–	b–
	MM	Mm	mm	M–	m–
7. Mouth					
8. Color	CC	Cc	cc	C–	c–
9. Tail	TT	Tt	tt	T–	t–

Remember: the allele for oval mouth (**M**) is dominant to the allele for slit mouth (**m**)

NewWorm Genetics		Two NewWorm Genotypes			
		NewWorm1		NewWorm3	
Body: Flat: BB or Bb Round: bb					
Mouth: Oval: ?? Slit: ??					
Head: Broad: ?? Medium: ?? Narrow: ??					
Rings: No Rings: RR or Rr Rings: rr					
Color: Green: CC Brown: Cc Black: cc					
Tail (Male): Pointed: TT or Tt Blunt: tt					
Tail (Female): Pointed: T- Blunt: t-					
Note: The Tail gene is in the X chromosome. The — indicates that the gene is not present in the Y-chromosome.					
Sex: Males: XX Females: XY					

MONOHYBRID INHERITANCE I

Use the NewWorm1 and NewWorm3 genotypes to answer these questions about babies they might produce.

Color

10a. Fill in the chart (Punnett square) to figure out possible genotypes (**CC**, **Cc**, **cc**) for a baby's color.

10b. Will any of NewWorm1 & NewWorm3's babies be **green**?

Definitely yes _____ Maybe _____

Definitely no _____

10c. What are the chances that one of NewWorm1 and NewWorm3's babies will be **brown**?

0 _____ 1/4 _____ 1/2 _____ 3/4 _____ 1/1 _____

Tail

11a. Fill in the chart (Punnett square) to figure out possible genotypes (**TT**, **Tt**, **tt**) for a baby's tail.

11b. Will any of NewWorm1 & NewWorm3's babies have a **pointed tail**?

Definitely yes _____ Maybe _____

Definitely no _____

11c. What are the chances that one of NewWorm1 and NewWorm3's babies will be **female** AND have a **pointed tail**?

0 _____ 1/4 _____ 1/2 _____ 3/4 _____ 1/1 _____

11d. If NewWorm1 and NewWorm3 have a **female** baby, what are the chances that it will have a **blunt tail**?

0 _____ 1/4 _____ 1/2 _____ 3/4 _____ 1/1 _____

NewWorm Genetics	Two NewWorm Genotypes
Body: Flat: BB or Bb Round: bb	NewWorm1 NewWorm3
Mouth: Oval: ?? Slit: ??	 
Head: Broad: ?? Medium: ?? Narrow: ??	 
Rings: No Rings: RR or Rr Rings: rr	 
Color: Green: CC Brown: Cc Black: cc	 
Tail (Male): Pointed: TT or Tt Blunt: tt	 
Tail (Female): Pointed: T- Blunt: t- Note: The Tail gene is in the X chromosome. The — indicates that the gene is not present in the Y-chromosome.	
Sex: Males: XX Females: XY	

DIHYBRID INHERITANCE

Use the NewWorm1 and NewWorm3 genotypes to answer these questions about babies they might produce.

Body and Rings

12a. Will any of NewWorm1 & NewWorm3's babies have a **flat body AND no rings**?

Definitely yes _____ Maybe _____ Definitely no _____

12b. What are the chances that one of NewWorm1 and NewWorm3's babies will have a **flat body AND rings**?

0 ___ 1/8 ___ 1/4 ___ 3/8 ___ 1/2 ___ 3/4 ___ 1/1 ___

OR impossible to tell from what's given ___

Color and Rings

13a. Will any of NewWorm1 & NewWorm3's babies have a **green body AND rings**?

Definitely yes _____ Maybe _____ Definitely no _____

13b. What are the chances that one of NewWorm1 and NewWorm3's babies will have a **black body AND rings**?

0 ___ 1/8 ___ 1/4 ___ 3/8 ___ 1/2 ___ 3/4 ___ 1/1 ___

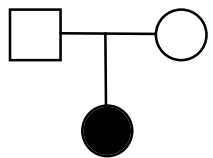
OR impossible to tell from what's given ___

PEDIGREE I: DOMINANCE RELATIONSHIPS

Consider two other NewWorm characteristics—Nostrils and Eyes.

- Each characteristic has two phenotypes as shown with the pedigree.
- Females are represented by circles and males are represented by squares.
- Decide what each pedigree says about the dominance relationship between each pair of phenotypes.

○ Large ● Small
□ nostrils ■ nostrils



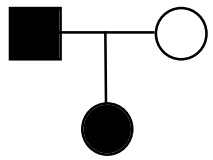
14. Having small nostrils is:

_____ definitely dominant

_____ definitely recessive

_____ impossible to tell from what's given

○ Red ● Yellow
□ eyes ■ eyes



15. Having yellow eyes is:

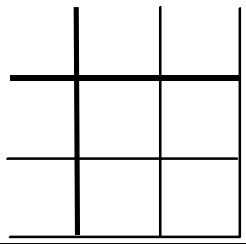
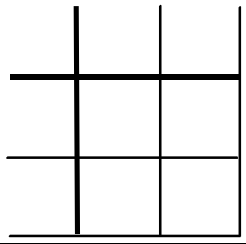
_____ definitely dominant

_____ definitely recessive

_____ impossible to tell from what's given

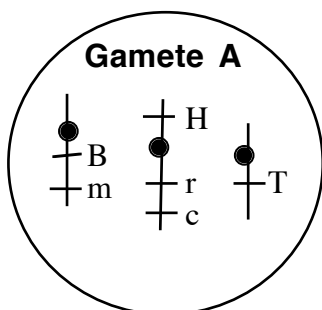
PUNNETT SQUARE

Label each Punnett square.

16. Write G in the spaces that represent gametes. Write O in the spaces that represent offspring.	
17. Write M in the spaces that represent the possible outcomes of meiosis. Write F in the spaces that represent the possible outcomes of fertilization.	

NewWorm Genetics		Two NewWorm Genotypes	
Body: Flat: BB or Bb Round: bb		NewWorm1	NewWorm3
Mouth: Oval: ?? Slit: ??			
Head: Broad: ?? Medium: ?? Narrow: ??			
Rings: No Rings: RR or Rr Rings: rr			
Color: Green: CC Brown: Cc Black: cc			
Tail (Male): Pointed: TT or Tt Blunt: tt			
Tail (Female): Pointed: T- Blunt: t-			
Note: The Tail gene is in the X chromosome. The — indicates that the gene is not present in the Y-chromosome. Sex: Males: XX Females: XY			

MEIOSIS: GAMETE A



18a. Was crossing over necessary for NewWorm1 to produce Gamete A?

Answer _____

18b. **If** you answered **yes**, circle the chromosome(s) in Gamete A that resulted from crossing over.

If you answered **no**, check here _____.

If you did not answer, do nothing.

MONOHYBRID INHERITANCE II: TEXTURE

Another inherited characteristic in the NewWorm is Texture. Both NewWorm1 and NewWorm3 have wrinkled skin. However when you mate them and produce 100 offspring, you find:

- 78 (38 males and 40 females) have wrinkled skin
- 22 (11 males and 11 females) have smooth skin

Remember: Males are XX and females are XY.

19a. There are two alleles for Texture. Is the relationship between the two alleles complete dominance or incomplete dominance?

Answer: _____

19b. **If** one of the Texture alleles is dominant, which one is it (wrinkled, smooth, OR neither)?

Answer: _____

19c. The gene for texture is autosomal.

What is it about the **offspring data** that indicates that the Texture gene is autosomal?

MONOHYBRID INHERITANCE II: EYELIDS

Another inherited characteristic in the NewWorm is Eyelids. Both NewWorm1 and NewWorm3 have clear eyelids. However when you mate them and produce 100 offspring, you find:

- 74 (51 males and 23 females) have clear eyelids
- 26 (0 males and 26 females) have cloudy eyelids

Remember: Males are XX and females are XY.

20a. There are two alleles for Eyelids. The relationship between the two alleles is complete dominance. What is it about the **offspring phenotypes** that indicates complete dominance?

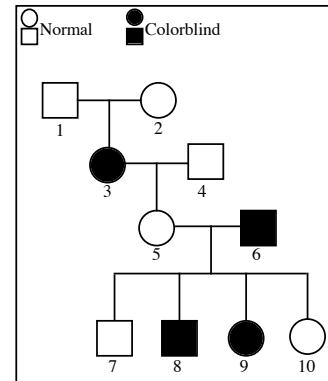
20b. The *clear* allele is dominant to the *cloudy* allele. The *cloudy* allele is recessive to the *clear* allele. What is it about the **offspring data** that shows you that *clear* is the dominant allele?

20c. Is the gene for Eyelids autosomal or X-linked? Answer: _____

PEDIGREE II: COLOR VISION–AUTOSOMAL OR X-LINKED?

Consider another NewWorm characteristic–Color Vision.

- Color Vision has two phenotypes as shown with the pedigree.
- Females are represented by circles and males are represented by squares.
- **Remember:** Males are XX and females are XY.
- Decide if the pedigree is consistent with Color Vision being autosomal or X-linked.



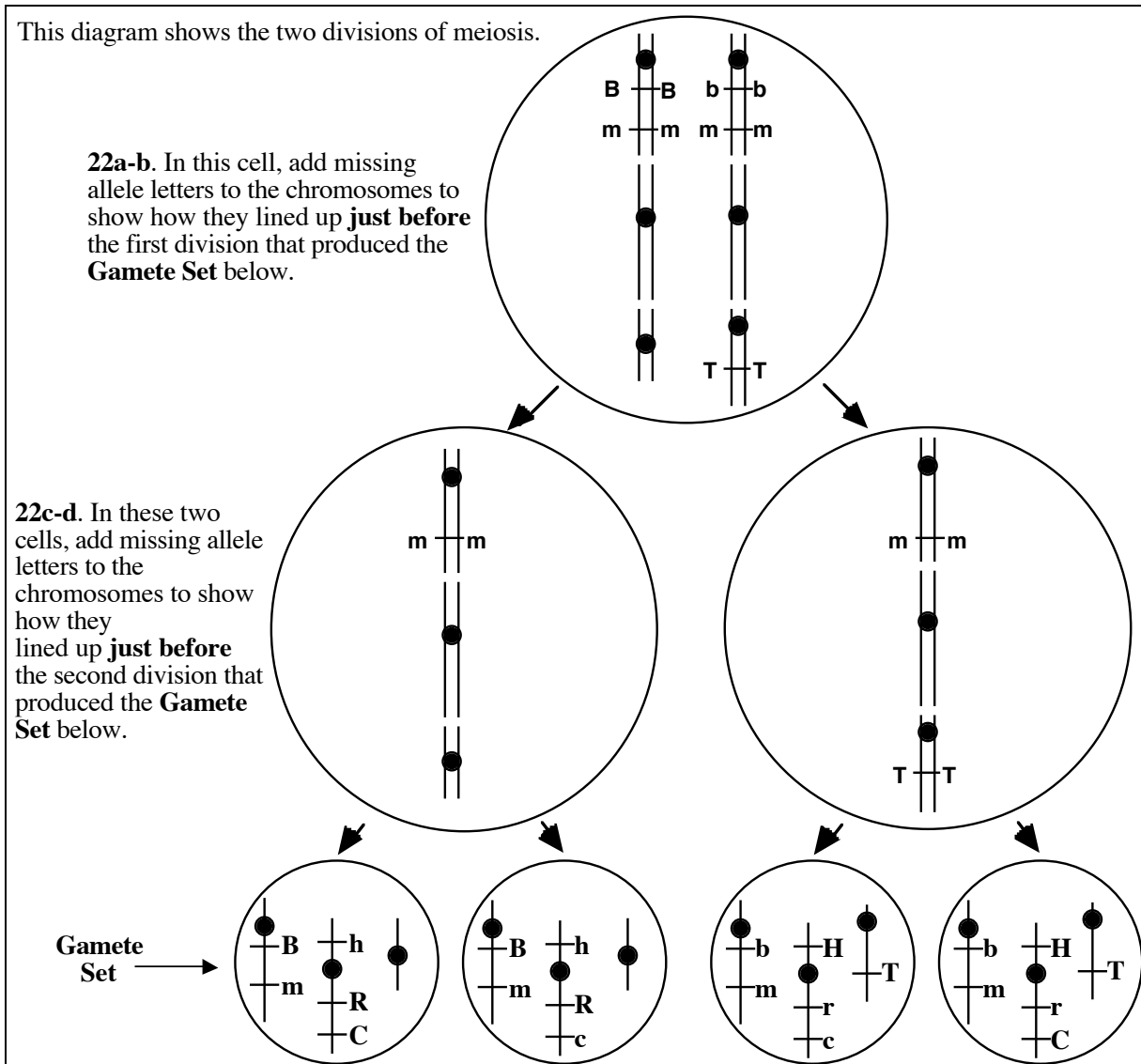
- 21a. Does the Color Vision gene **appear** to be autosomal or X-linked? Answer :
- 21b. Using words and/or diagrams, explain your answer to 21a (use the numbers below each circle or square to refer to particular individuals).
- 21c. Does this pedigree rule out the type of inheritance you did not choose? Answer :
- 21d. Using words and/or diagrams, explain your answer to 21c (use the numbers below each circle or square to refer to particular individuals).

This diagram shows the two divisions of meiosis.

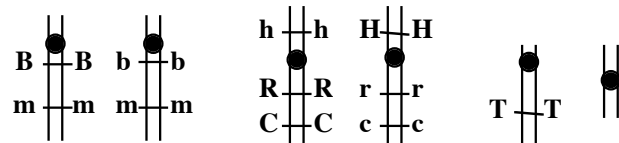
22a-b. In this cell, add missing allele letters to the chromosomes to show how they lined up **just before** the first division that produced the **Gamete Set** below.

22c-d. In these two cells, add missing allele letters to the chromosomes to show how they lined up **just before** the second division that produced the **Gamete Set** below.

Gamete Set →



The diagram to the right shows NewWorm1's chromosomes at the **beginning** of meiosis.



22e. Use **arrows** on the diagram to the right to show the exact location of any crossovers needed to produce the gamete set above.

NewWorm Genetics	Two NewWorm Genotypes
Body: Flat: BB or Bb Round: bb Mouth: Oval: ?? Slit: ?? Head: Broad: ?? Medium: ?? Narrow: ?? Rings: No Rings: RR or Rr Rings: rr Color: Green: CC Brown: Cc Black: cc Tail (Male): Pointed: TT or Tt Blunt: tt Tail (Female): Pointed: T- Blunt: t- Note: The Tail gene is in the X chromosome. The — indicates that the gene is not present in the Y-chromosome. Sex: Males: XX Females: XY	NewWorm1 NewWorm3

Probability I

- 23a. In a mating of NewWorm1 and NewWorm3, the chance of getting babies with **rings** is 1/2 or 50%.
- 23b. Draw a Punnett square that shows this. **Include genotypes and phenotypes.**
24. Explain and/or draw a diagram that shows how the way that chromosomes line up and separate during meiosis in NewWorm1 contributes to the 1/2 chance of getting worms with rings from this cross.

APPENDIX F

MULTIPLE CHOICE TEST

Multiple Choice Genetics Test

Mark through the letter of the answer you believe is correct. If you do not know the correct answer, but can eliminate one or more of the possibilities, then you should select the answer you think is most correct. However, because points are deducted from your score for each incorrect answer, you should not answer questions where you can't eliminate at least one incorrect response. In other words, guessing blindly won't increase your score.

- The first filial (F_1) generation is the result of
 - Cross-pollination among parents and the next generation.
 - Crosses between individuals of the parental generation.
 - Crosses between the offspring of a parental cross.
 - Self-fertilization between parental stock.
- To describe how traits can disappear and reappear in a certain pattern from generation to generation, Mendel proposed
 - The law of independent assortment
 - The law of segregation
 - The law of genotypes
 - That the F_2 generation will only produce purple flowers.
- In garden peas, a single gene controls stem length. The recessive allele (t) produces short stems when homozygous. The dominant allele (T) produces long stems. A short-stemmed plant is crossed with a heterozygous long-stemmed plant. Which of the following represents the expected phenotypes of the offspring and the ratio in which they will occur?
 - 3 long-stemmed plants: 1 short-stemmed plant
 - 1 long-stemmed plants: 1 short-stemmed plant
 - 1 long-stemmed plants: 3 short-stemmed plant
 - Long-stemmed plants only
 - Short-stemmed plant only
- True-breeding pea plants always
 - are pollinated by hand.
 - produce offspring with either form of a trait.
 - produce offspring with only one form of a trait.
 - are heterozygous.
- In fruit flies, a gene for body length is X-linked and long body alleles (B) are dominant to short body alleles (b). What are the possible genotypes of a long-bodied fruit fly?
 - Bb, bb, b-
 - bb, BB, B-
 - b-, bb, B-
 - B-, BB, b-
 - BB, B-, Bb
- The X and Y chromosomes are called the
 - extra chromosomes
 - phenotypes
 - sex chromosomes
 - all of the above
- In *Drosophila*, normal wing is dominant to wingless. In an experiment, a normal-winged male whose father was wingless was crossed with a wingless female. What percentage of the offspring would be expected to have normal wings?
 - 0%
 - 25%
 - 50%
 - 75%
 - 100%

Questions 8 and 9

refer to the Punnett square shown to the right. In humans, having freckles (F) is dominant to not having freckles (f).

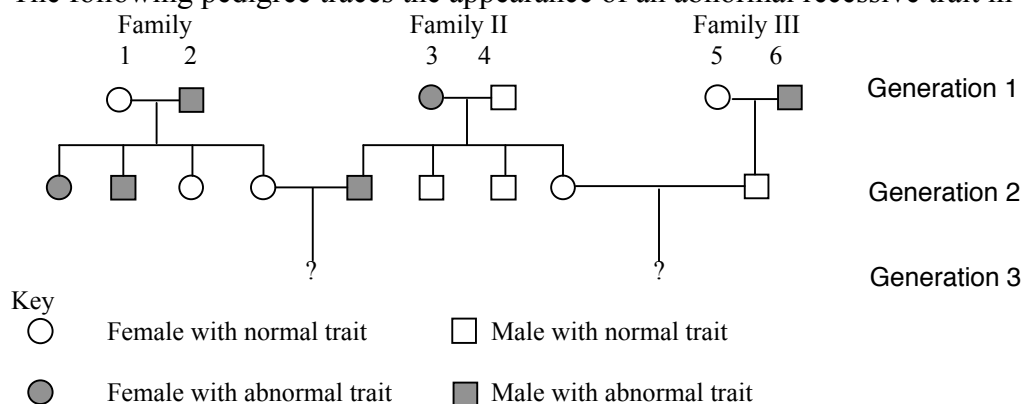
		Male Ff ♂	
		F	f
Female Ff ♀	F	1	2
	f	3	4

- The child who could be represented in box 1 in the Punnett square above would
 - be homozygous for freckles.
 - have an extra freckles chromosome.
 - be heterozygous for freckles.
 - not have freckles.

9. Which box in the Punnett square above represents a child who does not have freckles?
 (A) Box 1 (C) Box 3
 (B) Box 2 (D) Box 4
10. Cheetahs have been called the fastest animals on earth. They can achieve a burst of speed up to 110 km/hr. According to Darwin's theory of evolution, what is the best explanation for how such a capability might have evolved?
 (A) Fast-running cheetahs were more likely to survive and reproduce than those that ran more slowly.
 (B) Fast-running cheetahs were more likely to maintain constant body temperature than those that ran more slowly.
 (C) The more cheetahs ran, the stronger and faster they became.
 (D) Cheetahs show a preference for prey that run fast.
 (E) Individual cheetahs learned to associate fast running with survival.
11. Since the allele for color blindness is located on the X chromosome, color blindness
 (A) cannot be inherited
 (B) occurs only in adults
 (C) is sex-linked
 (D) none of the above
12. Gene technology is making it possible to correct genetic disorders by
 (A) identifying the carriers of genetic disorders and sterilizing them.
 (B) replacing copies of defective genes with copies of healthy ones.
 (C) using radiation treatments.
 (D) allowing parents to choose which sex they wish their children to be.

Question 13 refers to the diagram below.

The following pedigree traces the appearance of an abnormal recessive trait in several families.



13. Of the following, which provides evidence that the abnormal trait is NOT sex-linked
 (A) Some daughters of individuals 1 and 2 are normal for the trait.
 (B) Some sons of individuals 3 and 4 are normal for the trait.
 (C) Half the individuals in generation 1 have the trait.
 (D) Both sexes have the abnormal condition of the trait.
 (E) The trait appears in every generation.

APPENDIX G

TRANSCRIPTION KEY

=	latching
[	onset of overlapping speech
(.)	gap
(1.0)	length of silence in seconds
(())	transcriber's description
()	incomprehensible speech
:	prolongation of prior sound
a b c	indicates reading from quiz or answer explanation