5-12-2005

Transactive Discourse during Assessment Conversations on Science Learning

Homer Arthur Russell III

ACCEPTANCE

This dissertation, TRANSACTIVE DISCOURSE DURING ASSESSMENT CONVERSATIONS ON SCIENCE LEARNING, by HOMER A. RUSSELL III, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree Doctor of Philosophy in the College of Education, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chair, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty. The Dean of the College of Education concurs.

Dennis N. Thompson, Ph.D.
Committee Chair

Ann C. Kruger, Ph.D.
Committee Member

Karen Zabrucky, Ph.D.
Committee Member

John H. Neel, Ph.D.
Committee Member

Daniel T. Hickey, Ph.D.
Committee Member

Date

David B. Center, Ph.D.
Chair, Department of Educational Psychology and Special Education

Ronald P. Colarusso, Ed.D.
Dean

AUTHOR'S STATEMENT

By presenting this dissertation as partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education's director of graduate studies and research, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

_____

Signature of Author

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Homer A. Russell III
315 Vickers Drive, N. E.
Atlanta, GA 30307-1264

The director of this dissertation is:

Dr. Dennis N. Thompson
Department of Educational Psychology and Special Education
College of Education
Georgia State University
Atlanta, GA 30303-3083

VITA

Homer A. Russell III

ADDRESS: 315 Vickers Drive, N. E.
Atlanta, Georgia 30307-1264

EDUCATION:

Ph.D.	2005	Georgia State University Educational Psychology
M.S.	1997	Georgia State University Educational Psychology
B.S.	1984	Saint Mary's College Psychology

PROFESSIONAL SOCIETIES AND ORGANIZATIONS:

2004-Present	American Educational Research Association
2004-Present	Phi Lambda Theta

PRESENTATIONS AND PUBLICATIONS:

Fredrick, L., Bable, B., Russell, H. A., & Schafer N. (2004). *Behavior analytic approaches to enhancing engagement and measuring learning*. Poster presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.

Kruger, A. C., Schafer, N., & Russell, H. A., (2004). *A comparison of four different ways of coding engagement in videos of classroom activity*. Poster presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.

Russell, H. A., Kruger, A. C., Schafer, N. (2004). *Analysis of transactional discourse in learning*. Poster presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.

Schafer, N., Kruger, A. C., Russell, H. A. (2004). *Coaching video feedback for enhancing classroom assessment conversation and learning*. Poster presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.

Hickey, D. T., Kruger, A.C., Fredrick, L.D., Schafer, N.J., Russell, H.A., Bable, B., Hand, B., Michael, M., & Zuiker, S. (2003*). Looking at assessment and motivation and multiple perspectives: The GenScope assessment design experiment*. Paper presented at the 2003 meeting of the American Educational Research Association, Chicago, IL.

Ogden, P., Thompson, D., Russell, A., & Simons, C. (2003). The short- and long- term impact of supplemental instruction. *Journal of Developmental Education*, 26(3), 2-9.

Schafer, N. J., Hickey, D. T., Zuiker, S., Kruger, A. C., & Russell, H. A. (2003). *Using formative video feedback to facilitate classroom assessment conversation*. Paper presented at the 2003 meeting of the American Educational Research Association, Chicago, IL.

Russell, H. A. (1998). *A Consideration of multimedia instruction*. Atlanta, GA: Georgia State University. Available from http://education.gsu.edu/spehar/FOCUS/EdPsy/misc/Consider1.htm.

Russell, H. A., & Thompson, D. N. (1998). Enhancing reading comprehension in older adults. *Journal of College and Adult Reading and Learning*, 4(2), 29-34.

Russell, H. A., Beggs, R. J., May, G. L., & Oshima, T. C. (1997). *SPSS For Windows, Versions 7.5 and 8.0; A survival guide for EPRS 853 and 854, edition 3*. Available from http://education.gsu.edu/spehar/FOCUS/Spss.htm.

Hess, M. D., & Russell, H. A. (1988, July-August). Chemical warfare. *Infantry Magazine,* 38-40.

ABSTRACT

TRANSACTIVE DISCOURSE DURING ASSESSMENT CONVERSATIONS ON SCIENCE LEARNING
by
Homer A. Russell III

It has been argued that development of science knowledge is the result of social interaction and adoption of shared understandings between teachers and students. A part of understanding that process is determining how student reasoning develops in groups. Transactive discussion is a form of negotiation between group members as they interpret the meaning of their logical statements about a topic. More importantly, it is a form of discourse that often leads to cognitive change as a result of the interaction between group participants as they wrestle with their different perspectives in order to achieve a common understanding.

The research reported here was a correlational study designed to investigate the relationship between the various forms of transactive discussion and learning outcome performance seen in an investigation involving 24 students in a middle-SES high school located in southwest Atlanta, Georgia. Pretest and posttest measures of genetics reasoning, as well as curriculum content test data, were used in this study. Group discussion was captured on videotape and analyzed to determine whether transactional discussion was present and whether or not it had an effect on learning outcome measures.

Results of this study showed that participant use of transactive discussion played a role in development of reasoning abilities in the area of genetics. It is suggested that teachers should monitor classroom discourse for the presence of transactive discussion as such discourse plays a role in fostering performance outcomes.

TRANSACTIVE DISCOURSE DURING ASSESSMENT CONVERSATIONS ON SCIENCE LEARNING

by
H. A. Russell III

A Dissertation

Presented in Partial Fulfillment of Requirements for the
Degree of
Doctor of Philosophy
in
Educational Psychology
in
the Department of Educational Psychology and Special Education
in
the College of Education
Georgia State University

Atlanta, Georgia
2005

ACKNOWLEDGEMENTS

I am indebted to my dissertation committee members, Dennis Thompson, Ann Kruger, Karen Zabrucky, Dan Hickey, and John Neel, for their guidance, forbearance, and support. I am deeply indebted to Dennis Thompson for the advice, guidance, and mentoring he provided as my dissertation committee chair and as my doctoral program advisor.

I am especially indebted to my wife, Jane, for her loving support of my endeavors throughout my masters and doctoral programs. I also thank my family, and especially my parents, Russ and Linda Russell, for the support and encouragement of my efforts that only a loving family can provide.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance

CHAPTER 1

THE PROBLEM

Facilitating student participation in the domain of science has been a subject of debate across the past decade (National Academy of Sciences, 1995b). One result of that debate was publication of the National Science Education Standards (National Academy of Sciences, 1995a) in an effort to identify important aspects of a science education, and guarantee that students understand and can make informed decisions about science topics (National Academy of Sciences, 1995b). The development of science knowledge and reasoning ability does not occur in a vacuum, but is the result of social interaction and the development of shared understandings of science with teachers and peers in the science classroom (Driver, Asoko, Leach, Mortimer, & Scott, 1994). As part of that process, discerning how conversations within peer groups shape student reasoning was seen as key to developing more effective curricula (Jimenez-Aleixandre, Rodriguez, & Duschl, 1999).

The Problem

The assumption behind the research reported here was that a particular type of discourse known as transactive discussion has an effect on student outcome performance. The research reported here was a correlational study intended to address two general questions in order to help clarify the relationship between these factors. One general question the study addressed is whether transactive discussion was present in science

classroom peer collaborative activities. A second general question the study addressed is whether there was a relationship between transactive reasoning and outcome performance in science classrooms. Specific research questions will be addressed individually in Chapter 2 and will more fully explore the nature of the relationship between transactive discussion and performance on outcome performance.

The development of scientific thinking in peer groups is affected by a form of collaborative cognition known as transactive memory (Wegner, Giuliano, & Hertel, 1985). This form of thinking naturally results when participants work closely together in a peer collaborative environment and share a common understanding of ideas in an intellectual process of give and take (King, 1998). As a result, participants think in ways that they would not be able to do otherwise. This suggests that in many ways participants' thoughts are interconnected as they interact to encode, elaborate, label, store, and retrieve information.

As addressed in this study, transactive discussion is seen as a form of transactive memory. It is the interaction occurring between peers as they wrestle with different aspects of logical statements discussed in order to achieve a common understanding within a domain (Kruger, 1992). Transactive discussion was seen as well by Berkowitz and Gibbs (1983) as "reasoning that operates on the reasoning of another" (p.402). Transactive discussion can also be expressed in terms of the types of transactive dialogue behaviors, or transacts, that were seen occurring between discussants in a process Berkowitz and Gibbs labeled "transaction" (Berkowitz, 1980; Berkowitz & Gibbs, 1983). Moreover, Kruger and Tomasello (1986) noted that discussants "transact" when they participate in transactive discussions that engage a partner's reasoning. Importantly,

research has shown that transactive discussion contributes to improvement of outcome performance in scientific reasoning (Azmitia & Montgomery, 1993). These changes in outcome performance can be seen as a result of formative feedback (Torrance & Pryor, 1998) and discussants' responses to the opportunity to critically evaluate each others' ideas (Kruger, 1992) as they coordinate their conceptual understandings (Piaget & Inhelder, 1969).

The Setting

The study reported here was framed within the second year of a multi-year investigation known as the GenScope Assessment Project (Hickey, Kruger, Fredrick, Schafer, Zuiker, & Michael, 2004a). A primary goal of the project was to improve teaching and student performance. Another goal was to increase students' domain-specific classroom discourse to help improve student understanding and subsequent performance on both classroom assessments and external tests. Still another goal was to provide a point from which to reconcile differences between behaviorist, cognitivist, and situative/sociocultural views of knowing.

GenScope Assessment Project researchers pursued these goals in an iterative cycle of implementation-experimentation-evaluation to demonstrate the usefulness of design-experiment based research (Brown, 1992) in investigating challenging educational issues. Within the scope of the GenScope Assessment Project, one of the goals of the research reported here was to investigate the possibility of using transactive discourse to improve students' domain-specific discourse in the classroom.

Described elsewhere (Hickey et al., 2004a), GenScope Assessment Project researchers argued that their investigation was best understood in terms of a three by

three (3 X 3) framework of focal levels, models of domain knowledge, and cycles of design research. Focal levels were used to examine practice in the classroom. Each of the focal levels was seen as increasingly distant from the others. Additionally, GenScope Assessment Project researchers argued that each focal level reflected an increasingly formal model of domain knowledge.

Focal levels (close, proximal, and distal) were also interpreted by the three major perspectives within educational theory and practice (behaviorist/empiricist, cognitive/rationalist, and situative/pragmatist or sociocultural/sociohistoric) and their representations of domain knowledge. At the close level, situative/pragmatist or sociocultural/sociohistoric assumptions suggest that knowledge is represented in discourse and practices in a semi-formal manner. This representation in turn, is similar to curricular routines. At the proximal level, cognitive/rationalist assumptions suggest that knowledge is represented by formal comparisons to curriculum concepts. At the distal level, behaviorist/empiricist assumptions suggest that knowledge is represented by formal comparisons to targeted knowledge and skills.

GenScope Assessment Project researchers and others (Hickey, Zuiker, Taasoobshirazi, Schafer, and Michael, 2004b; Stein, 2001) noted that orchestrating productive classroom discourse is difficult. In an effort to foster productive classroom discourse, student interaction was structured around their interpretation and understanding of formative feedback assessments (Black & William, 1998). Teacher practices during implementation of the curriculum were also progressively fine-tuned in an effort to foster productive classroom discourse as well. In so doing, GenScope Assessment Project researchers used an iterative cycle of implementation-

experimentation-evaluation as suggested by notions of design-experiment research (Brown, 1992; Hickey et al., 2004a), which was intended to improve both classroom practice and research effectiveness.

In considering the assessment of change in student outcome performance, GenScope Assessment Project researchers argued that “activity-oriented” quizzes offered the best representation of domain knowledge at the close focal level. These quizzes were seen as observations of students' curricular activities that were interpreted using situative/pragmatist-sociohistoric or sociocultural perspectives. At the proximal level, exams were seen as "curricular-oriented." GenScope Assessment Project researchers argued that the best representation of domain knowledge at this level was seen in students' performance on questions that were interpreted using cognitive/rationalist perspectives. At the distal level, external standards-oriented assessments such as those mandated by programs like the *No Child Left Behind Act of 2001* were seen as "behavioral-oriented." GenScope Assessment Project researchers argued that the best representation of domain knowledge at this level was seen in performance on questions that are interpreted using behaviorist/empiricist perspectives.

GenScope Assessment Project researchers noted the National Research Council’s (2001) observation that the needs of classroom and external assessment were out of alignment and resulted in confusion in the classroom for teachers and students alike. They argued that the competing needs for internal and external assessment could be best met by focusing on assessment at an intermediate level. In doing so, the formative value of assessment at the close level could be compared to the performance on more distal student outcome measures. This also allowed the research reported here to focus on

discourse in the classroom with the expectation that changes in student outcome performance would be seen in assessments at the close, proximal, and distal levels.

The project was implemented across a three-year period using a 20-hour curriculum with several different teachers in different schools and classrooms being used for experiment and comparison. The research reported here describes results obtained in four classes taught by one teacher during the second year of the project and was aimed at enhancing student discourse to provide measurable increases in student outcome performance.

The GenScope Assessment Project consisted of the GenScope software program, GenScope curricular activities, three unit evaluations using formative assessment feedback and answer explanations, a final examination using formative assessment feedback and answer explanations, a "near-transfer" performance examination, and a "far-transfer" multiple-choice test (Hickey et al., 2004a).

The GenScope software program was designed to run on the generation of Macintosh computers common during the 1990s (Hickey, Kindfield, Horowitz, & Christie, 2003; Horowitz & Christie, 2000). The software was written to provide a program window representing any one of several different levels of biological organization appropriate for genetics instruction at the high school level. Biological organism information was graphically represented in each window, and the software provided tools that allowed students to interact with the organism at the level represented in the window. A key feature of the software was that it displayed the effects of student manipulation of the organism across levels; i.e., as a student made a change at one level, its effects were seen across all levels of the program. The software also modeled several

different organisms including humans, Labrador Retrievers, horses, and dragons. Most of the curricular activities, however, were designed around manipulations of dragons whose biological traits were simplified to facilitate student understanding (Hickey et al., 2003).

Fifteen GenScope curricular activities were designed to structure inquiry using a series of investigations that were completed by collaborating groups of students. Each investigation consisted of 1-3 pages of text and could be completed in a single class-period. Each investigation also involved use of laptop computers running the GenScope software. Additionally, the fifteen activities were spaced across four curricular units. At the completion of each curricular unit students were administered a quiz that was intended to help foster their understanding of the curricular materials they had just covered.

The first three quizzes focused solely upon the preceding curricular activities and were graded by the students during the following class period. Each quiz also prepared students for their subsequent performance on a challenging assessment that was administered before and after completion of the GenScope curriculum (known as the *NewWorm*). Additionally, each individual student graded the first three quizzes themselves during the next class session after administration of the quiz. When quizzes were returned to the students for grading, they also received supporting text-based formative feedback materials. One set of materials provided key instructional points for teachers implementing the GenScope curriculum. A second set of materials provided answers and explanations for each quiz question. The materials were written to force students to read the materials and interact with them to determine whether they had correctly answered a question. After grading their quizzes, students used another set of

structured materials to assess their understanding of a number of concepts covered by the quiz.

The fourth quiz administered at the end of the fourth week was written as a cumulative final examination (referred to as the final examination hereafter) and was graded by the teacher or GenScope Assessment Project researchers. This quiz also served as one of the dependent variables discussed in the research reported here. After the fourth quiz was returned, students received the same type of supporting formative feedback materials used in the previous 3 quizzes. They then went through the same processes of interaction with the previous quizzes as they sought to comprehend question explanations and assess their understanding of the concepts covered by the final examination.

Two additional assessments were conducted within the scope of the GenScope Assessment Projects. One was an assessment of "near-transfer" performance and reflected cognitive/rationalist perspectives. Known as the NewWorm assessment, it was text-based and written in a manner that approximated the GenScope curriculum, but using a similar organism which would be understandable to GenScope students and non-GenScope students alike. The NewWorm assessment was administered before and after the course of the GenScope curriculum and consisted of 25 short-answer questions that could be answered in about 40 minutes. GenScope Assessment Project researchers argued that the similarity between the NewWorm assessment and the GenScope curriculum meant that it functioned as a near-transfer, curriculum-based measure. They also argued that it also reflected higher-level knowledge structures that were constructed as students solved problems and engaged in discourse in the domain.

The second of the additional assessments evaluated "far-transfer" performance and reflected behaviorist/empiricist perspectives. Known as the SAT assessment, it was text-based, and consisted of multiple-choice questions taken from released versions of the SAT II-Biology and AP Biology tests. The SAT assessment was administered before and after the course of the GenScope curriculum, and consisted of nine multiple-choice questions that covered the entire range of difficulty represented in the released questions. Researchers chose assessment questions from an initial pool of 45 questions that had been ranked by difficulty and chosen to avoid similarity to the GenScope curriculum. From that pool of questions, GenScope researchers selected every fifth question. GenScope Assessment Project researchers argued the SAT assessment functioned as a "far-transfer measure" of outcome performance that represented the types of questions seen in external tests such as those students must pass to graduate. They also argued that the types of questions reflected associations between knowledge components developed by the students as they solved problems in the domain.

Scores for both the NewWorm assessment and the SAT assessment were scaled using Rasch measurement techniques (Acton, 2003) which are a type of one-parameter logistic model within item-response theory (Hambleton, 1985). The use of Rasch measurement techniques provided a means by which researchers could estimate an individual’s ability in relation to a number of related variables such as item difficulty, evaluator severity, and the challenge of the task. Use of this technique also allowed researchers to compare performance on the assessments between years as long as some assessment items remained in common between instruments.

GenScope Assessment Project researchers reported proficiency gains in performance on both the NewWorm and SAT assessments during the first year of the project. NewWorm assessment scores for the second year of the project, and the research reported here, reflected an average increase of 15.2, which was more than twice the average increase of 6.5 seen in the focal teacher's classes during year 1. Additionally, the year 2 increase in NewWorm assessment scores was six times greater than a gain of 2.5 seen in two non-GenScope classrooms used for comparison in the same school.

SAT assessment scores for the second year of the project showed a similar pattern of score improvement. SAT assessment scores reflected an average increase of 7.4, which was more than three times greater than the average increase of 2.1 seen in the focal teacher's classes during year 1. Additionally, the year 2 increase in SAT assessment scores was larger than the 5.7 gain seen in the two non-GenScope classrooms used for comparison, but not significantly so.

Variables of Interest

The initial independent variables were six different types of transactional behavior that were counted through analysis of the 32 videotapes made across four days. These six types of transactive behavior or transacts, represent a combination of the three types of transactive behavior analyzed by Kruger and Tomasello (1986) (as statement transacts, question transacts, and response transacts) and the type of possible interaction (self-oriented and other-oriented). Combined, the variables were: Other-Oriented Statement Transacts, Self-Oriented Statement Transacts, Other-Oriented Question Transacts, Self-Oriented Question Transacts, Other-Oriented Response Transacts, and Self-Oriented

Response Transacts. Additionally, the number of utterances containing no transactive behavior was noted as well.

Summative variables were created during the analysis by combining the different types of transactive behavior and non-transactive behavior in order to better understand the influence of different types and orientations of transactive discussion on outcome performance. These variables were Other-Oriented Transacts, Question Transacts, Transactive Utterances, and Total Utterances.

The study evaluated students' outcome performance as dependent variables. Outcome performance was assumed to be the result of students' processing of subject matter (Vermunt & Vermetten, 2004) encountered in the course of their participation in the GenScope Assessment Project. Five measures of outcome performance were selected: students' scores on an end-of-course final examination, NewWorm Pretest assessment, NewWorm Posttest assessment, SAT Pretest, and SAT Posttest.

The variables examined in this study are presented in Table 1. Table 1 presents each variable and its source.

Table 1

Variables of Interest

Variable Name	Composition
	Assessment Scores
Final Examination	Student final examination scores
NewWorm Posttest	NewWorm Posttest Scores
NewWorm Pretest	Student NewWorm pretest scores
SAT Posttest	Student SAT posttest scores
SAT Pretest	Student SAT pretest scores
	Frequency of Transactive Behavior
Other-Oriented Question	Count of Other-Oriented Question Transacts
Self-Oriented Question	Count of Self-Oriented Question Transacts
Other-Oriented Response	Count of Other-Oriented Response Transacts
Self-Oriented Response	Count of Self-Oriented Response Transacts
Other-Oriented Statement	Count of Other-Oriented Statement Transacts
Self-Oriented Statement	Count of Self-Oriented Statement Transacts
Non-Transactive Utterances	Count of Utterances without Transactive Utterances
	Summative Variables
Other-Oriented Transacts	Other-Oriented Question Transacts + Other-Oriented Statement Transacts + Other-Oriented Response Transacts
Question Transacts	Other-Oriented Question Transacts + Self-Oriented Question Transacts
Transactive Utterances	Other-Oriented Question Transacts + Self-Oriented Question Transacts + Other-Oriented Statement Transacts + Self-Oriented Statement Transacts + Other-Oriented Response Transacts + Self-Oriented Response Transacts
Total Utterances	Non-Transactive Utterances + Transactive Utterances

CHAPTER 2

REVIEW OF THE LITERATURE

The relationship of classroom discourse, peer collaboration, argumentation, transactive memory, and transactive discussion, to cognitive change is complex and little researched. There appears to be little literature uniting these factors. This review will illustrate the relationships between the factors considered in this study and seeks connections between them.

Each of the three major perspectives within educational theory and practice (behaviorist/empiricist, cognitive/rationalist, and situative/pragmatist or sociocultural/sociohistoric) would consider these factors according to their representative ways of knowing and learning. As noted by Greeno, Collis, and Resnick (1996), the nature of knowing and learning is different in each of these perspectives, and is framed in characteristic, and often complementary, ways.

They argued that the behaviorist/empiricist perspective (Greeno et al., 1996) represents knowing as organized interrelations, connections, and elements of expertise. This perspective also represents learning as a process by which these organized interrelations, connections, and elements of expertise are acquired. This perspective also allows that transfer is the degree to which the organized interrelations, connections, and elements of expertise can be applied in a new environment.

Greeno, et al. (1996) argued that the cognitive/rationalist perspective considers knowing as the building of understanding of concepts and theories in different domains, and understanding how to use cognitive skills. This perspective suggests that learning involves construction of concepts and growth of cognitive skills. Transfer is seen in this perspective as the degree to which one can apply understandings and skills in a new environment.

Greeno, et al. (1996) also argued that the situative/pragmatist or sociocultural/ sociohistoric perspective considers knowing as a function of how information is spread throughout cultures and individuals, their tools, and their participation in different communities of practice. This perspective suggests that learning is acquiring the ability to participate in the activities and practices of communities and cultures. Transfer is seen in this perspective as becoming attuned to the constraints and opportunities in a new environment.

The literature reviewed here in general adheres to a cognitive/rationalist perspective that is to a large degree Piagetian in orientation, although other perspectives are observed as well. Within the cognitive/rationalist perspective, Piaget (1964, 1971, 1995) focused on children's development of knowledge structures about the physical world around them as well as their social environment. Although Piaget did not focus on the social environment per se, he did recognize its importance in cognitive development (Piaget & Inhelder, 1969). He especially noted that social interaction offered opportunities for interaction between discussants in order to produce coordination of their understanding and cognitive skills (Piaget, 1964, 1995). Moreover, he also noted the importance of adult-child and child-child relationships in mediating cognitive change (De

Lisi & Golbeck, 1999). One often-studied environment providing opportunities for both adult-child and child-child interaction leading to cognitive change is that of classroom discourse.

Classroom Discourse

Discourse in the classroom has long been a topic of study. In his review of classroom discourse, Hartman (1996) noted studies of this topic as early as 1959, although there are doubtlessly others that were conducted even earlier still. Common to many studies in both the Piagetian and Vygotskian traditions has been the observation of the importance of classroom discourse for students in the development of understanding and knowledge in a number of domains (Cazden, 2001; Nathan & Knuth, 2003; National Council of Teachers of Mathematics, 1991; Nystrand & Gamoran, 1991; Palincsar, 1998; Webb & Palinscar, 1996).

Within the Piagetian perspective, classroom discourse was more often seen as an opportunity to facilitate child-child interaction, than as an opportunity to facilitate teacher-child interaction. Nevertheless, Piaget recognized the importance of teacher-child interactions as well (De Lisi & Golbeck, 1999; Forman & Kraker, 1985; Piaget, 1995; Piaget & Inhelder, 1969; Webb & Palinscar, 1996). In facilitating child-child interaction, Piaget believed their exchanges would provide opportunities for comparison, conflict, and reflection in order to provide opportunities for revision of their thoughts and beliefs (Berkowitz, Oser, & Althof, 1987; Landsmann, 1991; Piaget, 1964; Webb & Palinscar, 1996).

Within the Vygotskian perspective, classroom discourse has been seen as a social and cultural vehicle by which teachers assist children in the appropriation of ideas and

skills through participation in appropriate activities (Forman & McPhail, 1993; Rogoff, 1990; Torrance & Pryor, 1998). Unlike Piaget who did not focus on social interaction between children and adults (De Lisi & Golbeck, 1999; Forman & Cazden, 1994; Piaget & Inhelder, 1969; Webb & Palinscar, 1996), Vygotsky emphasized the importance of social interaction through discourse in fostering cognitive development (Forman, 1992; Hicks, 1996; Hogan & Tudge, 1999; Webb & Palinscar, 1996).

In considering the processes of classroom discourse, research from both traditions emphasized the importance of the teacher in orchestrating communication in the classroom (Brown, 1994; Cook, 2001a; Forman & Cazden, 1994; Forman, Larreamendy-Joerns, Stein, & Brown, 1998; Torrance & Pryor, 1998). Research from both traditions also emphasized that while classroom discourse may be effective in fostering cognitive development (Cazden, 2001; Crook, 1996; Forman et al., 1998; Nystrand & Gamoran, 1991; Palincsar, 1998), it is also difficult to organize successfully (Ball, 1996; Cazden & Beck, 2003; Chinn & Anderson, 1998; Osborne, Simon, & Erduran, 2002; Salomon & Globerson, 1989).

Additionally, research from both perspectives emphasized that the type of discourse occurring in the classroom was important (Cazden & Beck, 2003; Gee, 1989; Jimenez-Aleixandre et al, 1999; Nathan & Knuth, 2003; Nystrand & Gamoran, 1991; Nystrand, Gamoran, Zeiser, & Long, 2003; Orsolini & Pontecorvo, 1992; Williams & Butler, 1996) and affected cognitive outcomes. Two basic forms of discourse predominated in the classroom, although other forms may be possible. One predominant form of discourse in the classroom focused on the procedures, activities, structure, and social relations in the classroom, and though important, often centered around “doing the

lesson" (Jimenez-Aleixandre et al., 1999) and may have interfered with the academic aims of the classroom. The second predominant form of discourse in the classroom focused on academic discourse contributing to development of understanding and outcome performance, and centered around the substantive discourse associated with "doing the science" (Jimenez-Aleixandre et al., 1999).

Piagetian and Vygotskian perspectives acknowledge the importance of classroom discourse. However, as described below, rather than directly address how to increase classroom discourse, typical interventions representative of their perspectives often addressed practices that only indirectly affected classroom discourse. Other researchers (M. W. Berkowitz, personal communication, March 8, 2005) also noted the lack of efforts to directly transform classroom discourse as well.

Piagetian initiatives. Piagetian-inspired initiatives that addressed classroom discourse are limited and appeared to only approach the topic indirectly. One line of Piagetian influenced research investigated the effects of disagreement between discussants and was undertaken by a number of Piagetian influenced researchers (Berkowitz, 1980a, 1985; Berkowitz & Gibbs, 1985; Doise & Mugny, 1979; Doise, Mugny, & Perret-Clermont, 1976, Kruger & Tomasello, 1986; Miller & Brownell, 1975; Mugny & Doise, 1978; Perret-Clermont, 1980; Perret-Clermont & Schubauer-Leoni, 1981). Representative of this line of research was an early study that acknowledged Piagetian perspectives. Here Miller and Brownell (1975) investigated the nature of student discourse prior to agreement on the physical qualities of comparison objects. Their study involved 100 second grade students drawn from middle-class schools in Michigan. Students were paired together based upon their knowledge of the Piagetian

concept of conservation, the understanding that things remained the same despite changes in their form (Siegler & Alibali, 2005), with those understanding conservation being paired with a partner that did not understand conservation. Researchers interviewed student pairs after the students were exposed to conservation tasks. If the student pairs gave different answers to the researchers, they were asked to discuss the question and arrive at common answer with the result that students who understood conservation prevailed in student pair discussions because of their understanding of conservation. However, Miller and Brownell also found that students who understood conservation tasks engaged in more discourse than those students that did not understand conservation tasks.

Vygotskian initiatives. Vygotskian inspired initiatives often addressed discourse in the classroom (Anton, 1999; Cameron, 2002; Cazden & Beck, 2003; Hoel, 1997; Nathan & Knuth, 2003; Pontecorvo, 1993), but typically in an effort to understand discourse, rather than to use discourse itself to foster cognitive development. Representative of this type of research was Ge, Yamashiro, and Lee's (2000) investigation of students participation in collaborative activities over the Internet. In their research, they focused on the use of educational scaffolds, interventions intended to coach, involve and communicate processes, as a way to provide additional support to students in an Internet environment. Ge, Yamashiro, and Lee conducted a case study whose participants were undergraduate students, graduate students, and adult learners in a class taught at mid-sized university in the United States. Their case study observed actions before, during, and following student participation in a one-hour collaborative seminar conducted over

the Internet. They found that providing scaffolding activity before engaging in online or Internet activities facilitated collaboration and outcome performance.

However, some Vygotskian inspired initiatives have directly addressed classroom discourse and argument. Representative of this line of research was Orsolini and Pontecorvo's (1992) exploration of the discourse of children and teachers in the classroom in an effort to understand the mechanisms of interaction in their dialogue. Their study observed 12 low socioeconomic status preschool students (5-6 years of age) drawn from the same class outside Rome, Italy, during 11 small and large group discussions. They found that during agreement themed discussions, extended discussions by children were generally preceded by previous peer elaborations, or by the teacher's focused restatement of information. They also found that during disagreement themed discussions, peer dialogue tended to often be independent of the teacher's perspective and often argumentative. However, they also noted that peer dialogue approached the teacher's instructional aims during disagreement themed discussions.

Peer Collaboration

Like classroom discourse, the importance of peer collaboration has long been noted. Webb and Palincsar's (1996) review of classroom group processes noted that collaboration is the process of converging on shared understandings of meaning, ideas, and experience among discussants. Roschelle's (1992) review of collaboration further suggested that the Piagetian and sociocultural traditions recognized collaboration, but differed in their conception of how collaboration occurred. The sociocultural perspective viewed collaboration as a process of scaffolding by a more capable peer and appropriation of those concepts by the less capable peer (Forman, 1992; Forman &

Cazden, 1994; Forman & Kraker, 1985). In contrast, the Piagetian perspective focused on conflict between peers (Doise & Mugny, 1979; Doise et al., 1976; Perret-Clermont, 1980; Perret-Clermont & Schubauer-Leoni, 1981).

Furthermore, Piaget (1964, 1971, 1995), observed that social interaction was integral to cognitive development. Moreover, he also believed that the experience of collaboration changed the way in which children reasoned by forcing them to note and adjust their internal discourse and reasoning to meet that of their associates. In that process of adjustment, children learned rules and ways of thinking that approached and accommodated the alternate viewpoints offered by others (Piaget, 1964). Later researchers have shown that peer collaboration had a positive effect on learning outcomes in a number of settings, four of which are described here.

Math reasoning. In a two-year examination of the effect of peer collaboration on the development of spatial and mathematical reasoning, Phelps and Damon (1989) studied 152 fourth graders. They found that children in peer collaborative groups learned significantly more than children in control groups and suggested that peer collaboration fostered the development of reasoning skills. In a later study of the effect of peer collaboration on the participation and learning outcomes of low-achieving math students, Gabriele and Montecinos (2001) made a similar observation in finding that peer collaboration supported the development of learning outcomes as well.

More recently, Vauras, Ilskala, Kajamies, Kinnunen, and Lehtinen (2003) examined the effect of peer collaboration on mutual mediation of the development of learning outcomes and problem solving in a gaming environment that taught mathematical problem-analysis. In their study, eight high and low achieving students

participated together in an instructional game across an eight week period. They found that high ability students who actively participated in peer collaboration experienced improvements in their mastery of skills involved in solving mathematical problems. Additionally, they found that transactive discussion had a positive effect on students' performance.

Moral development. In an early study acknowledging Piagetian perspectives, Berkowitz, Gibbs, & Broughton (1980) investigated the effect when peer collaborators with different levels of moral development discussed moral dilemmas. Their study involved 41 pairs of undergraduate psychology students who discussed investigator-selected dilemmas across a two-month period. They found that discussion among collaborating partners was effective in providing changes in moral development when disparity in level of moral development was small, but was less effective when disparities were large.

In a more recent investigation that compared peer-child and adult-child collaborative groups, Kruger (1992) studied 24 child-child dyads and 24 adult-child dyads who collaborated in a discussion of moral dilemmas. She found that reasoning about moral dilemmas differed between child-child and adult-child pairings, with the child-child interactions being more spontaneously produced and active than the adult-child interactions, and her results supported the importance of peer collaboration in the development of moral judgment.

In a follow-up investigation, Kruger (1993) studied the nature of cognitive change during peer collaboration. The subjects were the same as in her earlier investigation mentioned above. However, in this study she found that peer collaboration was most

likely to result in cognitive change if suggested solutions to moral dilemmas were rejected. Other researchers also reported the importance of peer collaboration on moral development as well (Berkowitz & Gibbs, 1983; Damon & Killen, 1982; Tomasello, Kruger, & Ratner, 1993).

Scientific reasoning. A large body of research has described the positive effect of peer collaboration on the development of scientific reasoning. Thorley and Treagust (1987) examined conflict in peer collaboration and its effect on conceptual change in understanding physics. In their study, 21 college-age peer dyads collaborated in a discussion of electrical circuits and mechanics. Dyad members were paired to ensure inter-student conflict by selecting one member who held an accurate view of the discussed topic and one member who held an inaccurate view of the discussed topic. Results of the study demonstrated that interpersonal conflict and argument in peer collaboration were an important part of conceptual change.

Scardamalia and Bereiter (1992), studied the differences between text-based and knowledge-based questions asked by children. In their investigation, they conducted two separate studies with students from two different classes composed of both fifth and sixth grade students. In that investigation, Scardamalia and Bereiter noted that as children became more experienced with the study protocols, they moved from asking questions that helped them understand problems to asking questions that extended their knowledge of what they already knew. Scardamalia and Bereiter also found that in order for the students to ask productive questions, the students had to be members of a community where questioning was valued and where the students felt their contributions were appreciated.

In a later study, Lonning (1993) evaluated the effect of peer collaboration on conceptual change in a tenth grade general science class. In his study of 36 at-risk students, he found that peer collaboration fostered conceptual change as measured by the correct use of scientific terms.

Azmitia and Montgomery (1993) observed the effect friendship between collaborating peers had on the development of scientific reasoning. In their study, 72 fifth graders participated in a year-long examination of friendships in collaboration. They found that collaborating with friends resulted in the development of higher levels of scientific reasoning. Interestingly, they also found that it was especially important to work with a friend if a problem was difficult.

Okada and Simon (1997) examined collaboration in scientific problem solving and genetics. In their study, 27 undergraduate students participated in scientific problem solving either alone, or in collaboration with a friend. They found that the performance of collaborating peers was superior to that of individuals working alone. They also found that collaborating peers produced significantly more explanatory activities than individuals who worked alone.

Tao and Gunstone (1999) investigated the effect that the use of a computer had on collaboration in a physics classroom. In their study, 26 students were organized as 13 dyads who participated in a 10-week block of physics instruction. They found that conceptual change occurred when computers were used to support collaborative activities. More importantly, they found that the co-construction of ideas, argumentation, and peer conflict, as well as the opportunity for reflection, were key components that supported conceptual change during peer collaboration.

Leinonen, Virtanen, Hakkarainen, and Kligyte (2002, Jauary) researched collaboration in the use of knowledge-building discourse. In a pilot study, seven graduate students participated in a computer-mediated software environment which facilitated collaboration. They found that during initial collaborative events, the students were unfamiliar with study methods and expectations. However, as the students became more familiar with the process, they increased the amount and quality of their collaborative engagement.

More recently, Hakkarainen (2003) examined the effect of progressive-inquiry in a collaborative environment. In his study, he observed 145 students in two fifth and sixth grade classrooms across a three year period. He found that the quantity and quality of questioning among peers increased as a result of prolonged engagement in a supportive environment.

Still more recently, Ding (2003) examined the nature of collaborative scientific discovery. In his study, 60 undergraduate science majors were separated as paired-subjects on the basis of friendship. He found that participation in explanatory activities facilitated the generation of hypotheses that lead to cognitive change and scientific discovery. He also found that argument and disagreement between partners promoted, but did not guarantee cognitive change. Rather, the key issue noted by Ding (2003) was whether or not subjects noted their disagreement, and then adopted new and different strategies which subsequently lead to successful performance. Additionally, he found that cognitive change was most likely to occur when subjects accounted for the phenomena they observed. Moreover, he found that the process of accessing information though collaboration and reciprocal intellectual interaction stimulated new perspectives which

led to additional collaboration and reciprocal intellectual interaction. Other researchers reported the importance of peer collaboration on the development of scientific reasoning as well (Hakkarainen & Järvelä, 2002; Howe, Tolmie, & Rodgers, 1990, 1992; Messer, Joiner, Loveridge, Light, & Littleton, 1993; Teasley, 1995; Teasley & Roschelle, 1993; Tolmie, Howe, Mackenzie, & Greer, 1993; Tudge & Winterhoff, 1993).

Musical composition. Miell and MacDonald (2000) examined social processes that occurred when students collaborated in musical composition. In their study, 40 middle school children, 11 – 12 years of age, were separated into two groups of 20 paired students. In the first group, each pair consisted of friends. In the second group, the members were not previously friends. Miell and MacDonald found that friendship affected the collaborative process in a number of ways. One of the most important effects was that the amount of transactive discussion was higher among friends than it was among those who were not friends. Additionally, they also found that the amount of transactive discussion affected the quality of musical arrangements, with higher amounts of transactive discussion being associated with better musical compositions. Azmitia and Montgomery (1993) made a similar observation in their study of the development of scientific reasoning as well.

Argumentation

Duschl and Gitomer (1997) argued that one of the goals of an education in science is to develop the ability to reason about science claims. Piaget (1964; Piaget, 1971, 1995) argued that social interactions are needed to facilitate the development of individuals' skills in reasoning. Driver, Newton, and Osborne (2000) noted the importance of social interactions and made the case that scientific knowledge is socially constructed from

what we know of the real-world. They also stated that argumentation is socially situated and should have a central role in science education as it supported the development of an understanding of scientific knowledge and how it is constructed. Their position on argumentation followed from the earlier 1958 research of Toulmin that identified the elements that must be included to support an argument as valid: the *claim* one wishes to defend, the *data* upon which one bases the claim, the *warrants* that are the reasons a claim is supported by the data, a *qualifier* that may modify the nature of a warrant, and *backings* that provide the reasons a warrant should be accepted as authoritative. Knowledge of the elements of argumentation was not seen as enough to support reasoning. Also important was the environment and the discourse that supports argument.

Indeed, Kuhn (1991, 1993) specifically noted the importance of social discourse in argumentation and more explicitly, suggested that it offered a way to check our thoughts against those of others that might differ from or contrast with our own (Kuhn, 1991). In a 1992 investigation, Kuhn examined argumentive reasoning in a population of 160 subjects that included ninth-grade students, young adults, middle-aged adults, and older adults in their 60s. She found that argumentive reasoning was related to the level of one's education and argued that by about the ninth grade the development of one's argumentive skills was complete and changed little afterwards. She also found that only about 40 percent of the subjects used appropriate evidence in composing their arguments and that in general, adults, as well as children, had difficulty in developing arguments that were supported with appropriate evidence.

In another investigation, Kuhn, Black, Keselman, and Kaplan (2000) studied the development of cognitive skills associated with inquiry curricula in 42 sixth, seventh, and

eighth grade students. In that investigation, they found that the students had difficulty ordering and interpreting evidence. In a similar investigation, Felton and Kuhn (2001) researched the development of argumentive discourse skill in a study involving 31 college students and 33 seventh and eighth graders. They found that young teens had difficulty coordinating the evidence associated with argumentive discourse. Notably, in general the younger students did not seem to understand many of the goals of argumentive discourse, and appeared to be more interested in the production of the associated dialogue than in the purpose of the dialogue.

In different study, Kuhn, Shaw, and Felton (1997) conducted two investigations to investigate the effect of thinking and reasoning when subjects were involved in dyadic discussion. In the first investigation subjects were 49 seventh and eighth grade students and 44 adults. Only the students were considered in the second investigation. In both studies, each age group was separated into experimental and control groups and participated in dyadic discussions for a five week period. Overall, Kuhn and her associates found that, similar to other studies, the subjects failed to appropriately use evidence to support their arguments. They also found however, that as a result of collaborating in dyadic discussion, both adolescents and adults experienced conceptual change, with the adolescents being almost twice as likely to do so. More importantly, Kuhn and her associates found that sustained engagement in dyadic dialogue facilitated the development of thinking and reasoning in both age groups. They attributed this effect to social transmission of new knowledge as a result of sustained communication between study participants.

Transactive Memory

Lehtinen (2003) noted that during peer collaboration there appeared to be a very special form of communication that occurred between partners. Very often this special type of communication was found operating in the classroom as students collaborated in a reciprocal and interdependent manner, and was found to facilitate development of their learning outcomes (King, 1998). Termed *transactive memory* (Wegner et al., 1985), it was seen as a direct result of the cognitive interdependence that developed because of the communication between collaborating members (Ding, 2003).

Recognized as a form of shared memory (Wegner, Raymond, & Erber, 1991), transactive memory was seen as a process that allowed participants to think in ways that they would not have been able had they not participated in the relationship (Levine & Resnick, 1993; Moreland, Argote, & Krishnan, 1996). Moreover, as demonstrated in an investigation by Liang, Moreland, and Levine (1995) of the effect of transactive memory on group and individual performance in 90 undergraduate students, transactive memory fostered improvements in group achievement.

Wegner et al. (1985) described transactive memory as consisting of two components. One component was the cumulative task-relevant knowledge possessed by each collaborating member. A second component was the knowledge transactions or transactive processes that occurred in the course of communication between group members as they collaborated on a task (Wegner, 1987; Wegner et al., 1985). Arguably, two of the most important processes that occurred between group members were *transactive encoding* and *transactive decoding*. Together, they played a key role in the integration of individual and collective group knowledge about a collaborative project

(Wegner, 1987). In the process of transactive encoding and decoding, participants discussed information pertinent to their task, typically by negotiating the shape the information would take. In this same process, participants also recognized the expertise of individual members as information was translated into a form the group could use. The process of retrieval was similar in that information was recalled by a member as its importance, manner and means of presentation to the group were negotiated within the group.

Transactive Discussion

Berkowitz and Gibbs (1983) defined *transactive discussion* as "reasoning that operates on the reasoning of another" (p.402) and saw it as part of the communication process that occurred between collaborating partners. Berkowitz et al. (1987) specifically noted that transactive discussion was based upon the Piagetian assumption that cognitive development is the result of efforts to resolve differences between incompatible positions held by discussants. They also noted that transactive discussion was an advanced form of argumentation which although present in childhood, becomes more frequent as one gets older.

Transactive discussion can be seen as one of the processes of transactive memory cited by Wegner et al. (1985). As later argued by Wegner (1987), transactive encoding was seen as taking place when collaborative group members discussed information in a process of implicit negotiation about how data was to be understood and subsequently used by the group. This process of the negotiation of meaning appeared to be analogous to the process of transactive discussion and served to clarify the understanding of group members about a topic in a similar manner.

Berkowitz and Gibbs (1983) also defined 18 types of transactive behavior seen during transactive discussion which they termed *transacts*. They believed these transacts were the processes that occurred during conversations and which subsequently promoted developmental change. Moreover, they emphasized the importance of the use of transacts between discussants in labeling the process as one of "transaction." In a similar sense, Kruger and Tomasello (1986) noted that when one "transacts," they operated on the reasoning of a partner. Kruger (1992) later also affirmed the importance of transaction in her observation that the process of transaction on several perspectives predicted cognitive change.

In their investigation of transactive discussions between groups of peers, and peers with adults, Kruger and Tomasello (1986) considered three different types of transactive behavior: *transactive statements*, *transactive questions*, and *transactive responses*. They defined transactive statements as voluntarily expressed restatements, improvements, evaluations, or elaborations of ideas. They defined transactive questions as voluntarily expressed requests for amplification, validation, or explanation of ideas. They defined transactive responses as voluntarily expressed amplification, validation, or explanation of ideas in response to a transactive question. In a later study Kruger (1992) further emphasized that it was the type of transactive behavior that was important, rather than just the quantity of transactive behavior.

King (1997) described a model for transactive peer tutoring. A key component of her approach was the use of transactive questions to scaffold participants' thinking. In emphasizing the use of transactive questions and questioning to facilitate development of learning outcomes, she continued the course of her earlier research which previously

emphasized the importance of the use of questions in preference to other methods (King, 1994, 1995; King & Rosenshine, 1993).

Kruger and Tomasello (1986) also considered how partners interacted with each other's ideas and defined this as orientation. They defined voluntarily expressed observations about a partner's ideas as *other-oriented*. They defined voluntarily expressed observations about one's own ideas as *self-oriented*. Kruger (1993) later found in her study of the nature of cognitive change during peer collaboration that the orientation of partners' transactive discussion had an effect on learning outcomes. Self-oriented transactive discussion appeared to be negatively correlated with cognitive gain and learning outcomes, while other-oriented transactive discussion was often seen as the basis for rejection of suggested solutions and cognitive gain.

Azmitia and Montgomery (1993) made a similar observation in their study of the effect of friendship between collaborating peers on the development of scientific reasoning. In that study, they found that there was a positive correlation between both self-oriented and other-oriented transactive discussion, and cognitive change. They also found that friends produced more other-oriented transactive behavior, which induced peer conflicts and facilitated problem solution.

Additionally, transactive discussion appears to have the greatest effect on learning outcomes when it operates across multiple viewpoints, both those that are accepted and most importantly, those that are ultimately rejected, and seems most likely to have occurred when partners disagreed (Azmitia & Montgomery, 1993). However, Kruger (1993) observed that it was not enough to have multiple points of view. Instead, she

argued that it was the critical examination of the various points of view that fostered cognitive change.

Summary

The relationship between classroom discourse, peer collaboration, argumentation, transactive memory, and transactive discussion is complex, with each of these topics having their own literatures. There is a robust literature addressing classroom discourse and its importance in the development of student learning outcomes. The reviewed literature emphasized the importance of classroom discourse as a mediator of cognitive performance outcomes and learning in the school setting, principally by fostering child-child and teacher-child interactions. However, the literature also emphasized that productive discourse is difficult to orchestrate in the classroom. Additionally, the reviewed literature showed that research tended to examine classroom discourse as a high-level process, rather than as an agent that could be manipulated directly to facilitate cognitive development.

There is an extensive literature detailing the effect of peer collaboration on the development of learning outcomes. This review examined a small portion of the literature that detailed the positive effect of peer collaboration on learning outcomes in the areas of math reasoning, moral development, scientific reasoning, and musical composition. The reviewed literature shows that the peer collaborative process fostered cognitive change. Additionally, argumentation, transactive memory, and transactive discussion have been shown to be associated with, if not a part of the process of peer collaboration.

The literature also shows that argumentation fosters cognitive change through several different means. One way argumentation fosters cognitive change is by providing

an avenue for social discourse that allows participants to check their thoughts against the thoughts of others. A second way that argumentation fosters cognitive change is by providing an opportunity for sustained engagement about a topic.

Transactive memory and transactive discussion have been shown to be associated with cognitive change in collaborative peer groups. Transactive memory systems were shown to facilitate consideration of alternate perspectives and cognitive change by providing a common understanding of information held by members of the collaborative group.

Transactive discussion was shown to foster cognitive change by facilitating the interaction of peers and the examination of alternate points of view. Individually and in various combinations, peer collaboration, argumentation, transactive memory, and transactive discussion have been shown in the literature to have an effect on cognitive change. The literature reviewed here was chosen to illustrate the effect of each factor on cognitive change and to emphasize the role played by transactive discussion.

Research Questions

Research Question 1: What is the amount, type, and orientation of transacts that students generate overall during their feedback discussions?

The first research question was intended to provide a basic description of the nature of transactive discussion seen during the investigation. The study used data gathered through the 2002 instantiation of the GenScope Assessment Project (Hickey, 1999). That project videotaped student triads in an assessment of student engagement and participation in a ninth grade biology classroom studying genetics. The videotapes were coded for transactional behavior in a manner consistent with Kruger's (1992; 1993)

studies. More specifically, the videotapes were made during formative feedback (Black, 1998) discussions conducted during the class, following administration of three content tests and a final examination.

As noted earlier, research has shown that the amount of transactive discussion affects performance on learning outcomes (Kruger & Tomasello, 1986; Miell & MacDonald, 2000). Other research suggests that it is the type of transactive behavior that affects performance on learning outcomes (Berkowitiz & Gibbs, 1983; Kruger, 1993; Kruger & Tomasello, 1986). Additionally, research has shown that the orientation of peers' transactive discussion has an effect on learning outcomes (Azmitia & Montgomery, 1993; Kruger, 1993). Measuring the amount, type, and orientation of transactive behavior uttered by the videotaped students is a first step in understanding the effect of transactive discussion on learning outcomes.

Research Question 2: Do students improve from pretest to posttest on scores of the SAT and NewWorm assessments?

The second research question was intended to investigate whether or not there was an improvement from pretest to posttest on measures of far-transfer and near-transfer (Hickey, Kruger, Fredrick, Schafer, & Kindfield, 2002). Pretest and posttest comparison data was gathered using two assessments. The first was a SAT assessment which consisted of items selected from released SAT tests. The second was the NewWorm assessment which was created during an earlier investigation as a measure of subjects' skill in thinking about genetics (Hickey & McCaslin, 2000). This question investigates whether or not there was a change in performance on these assessments of far-transfer and near-transfer.

Research Question 3: Does the total number of transacts increase over time?

The third research question was intended to investigate whether or not there was a change in the amount of transactive behavior seen over time. The expectation of an increase in communicative behaviors is not unwarranted. Scardamalia and Bereiter (1992) noted that as children became more experienced in an instructional environment they moved from asking questions that helped them understand problems, to asking questions that extended their knowledge of what they already knew. Leinonen et al. (2002, January) made a similar observation in their investigation and found that the quality and quantity of collaboration increased over time as task engagement grew longer. Additionally, Hakkarainen (2003) observed that the quality of questioning improved as when engagement was prolonged in a supportive classroom. These points are similar to and support Ding's (2003) observation, noted previously, that the process of accessing information through transactive processes led to additional transaction.

Research Question 4: W*ill other-oriented transacts and question transacts become more frequent across the four feedback assessments examined in the GenScope research?*

The fourth research question was intended to investigate the incidence of other-oriented and question transactive behavior in a more focused manner than research question three. Given that transacts increased over time as asked in the third research question, this question asks if there is a difference in orientation and type of transactive behavior that changes over time. The importance of orientation of transactive behavior is shown in Kruger's (1993) observation that other-oriented transactive discussion often led to cognitive gain. The importance of the type of transactive behavior is shown in Scardamalia and Bereiter's (1992) observation that questions were used more frequently

as students became familiar with their use. The importance of the type of transactive behavior is also shown in Hakkarainen's (2003) research that reported the quality of questioning improved over time. More specifically, the importance of the type of transactive behavior is shown in King's (1997) observation that students learned to use question transacts more effectively over time.

Research Question 5: Is there a relationship between the number of other-oriented transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?

The fifth research question proceeded from the first research question and was intended to determine whether other-oriented transactive behavior had an effect on performance as measured by the SAT and NewWorm assessment posttests, and the final examination. As previously noted, Kruger (1993) showed the importance of orientation of transactive behavior in observing that other-oriented transactive discussion often led to cognitive gain. Also as noted earlier, pretest and posttest comparison data was gathered using two assessments; the SAT assessment, and the NewWorm assessment. Additionally, three content tests and a final examination were administered during the assessment project and served as a measure of content knowledge. This question investigates whether or not there was a change in performance on these assessments as a result of the use of other-oriented transactive behavior.

Research Question 6: Is there a relationship between the number of question transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?

The sixth research question proceeded from the first research question and was intended to determine whether question transactive behavior had an effect on performance as measured by the SAT and NewWorm assessment posttests, and the final examination. As previously noted, King (1997) showed that students learned to use question transacts more effectively over time. Additionally Scardamalia and Bereiter (1992), and Hakkarainen (2003) observed the importance of questioning in developing learning outcomes.

Also as noted earlier, pretest and posttest comparison data was gathered using two assessments; the SAT assessment, and the NewWorm assessment. Additionally, three content tests and a final examination were administered during the assessment project and served as a measure of content knowledge. This question investigates whether or not there was a change in performance on these assessments as a result of the use of transactive questions.

Research Question 7: Is there a relationship between the total number of transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?

The seventh research question proceeded from the first research question and was intended to determine whether the total of all transactive behavior had an effect on performance as measured by the SAT and NewWorm assessment posttests, and the final examination. Miell and MacDonald (2000) found that the amount of transactive discussion affected performance, with higher amounts of transactive discussion being associated with improved performance. However, Kruger (1992) suggested that it was the

type of transactive behavior that was important, rather than the quantity of transactive behavior.

Also as noted earlier, pretest and posttest comparison data was gathered using two assessments; the SAT assessment, and the NewWorm assessment. Additionally, three content tests and a final examination were administered during the assessment project and served as a measure of content knowledge. This question investigates whether or not there was a change in performance on these assessments as a result of the total of all transactive behavior as noted Miell and MacDonald (2000).

CHAPTER 3

METHODOLOGY

Participants

Participants were twenty-four ninth grade students (12 female, 12 male) in a middle class, predominantly African-American high school located in southwest Atlanta, Georgia, who took part in a specially designed program on genetics intended to support the GenScope Assessment Project. All students, except for one, were of African-American heritage and reflected the 99.5% African-American population of the school. Additionally, about 30% of students in the school were qualified for a federal lunch subsidy. However, unknown is how many or if any of the students participating in the study also participated the lunch subsidy.

Measures

Student knowledge of genetics domain information was evaluated through two assessments. One assessment, referred to in the study as the SAT assessment, consisted of items selected from released SAT tests. The second was the NewWorm assessment which was created during an earlier investigation as a measure of students' genetics reasoning. Both measures assessed changes in genetics knowledge and reasoning, and were administered pretest and posttest. Performance on these assessments was evaluated as a part of a larger study (Hickey et al., 2003) using Rasch measurement techniques (Acton, 2003) which are a type of one-parameter logistic model within item-response theory

(Hambleton, 1985). The use of Rasch measurement techniques provides a means by which an individual's ability can be estimated in relation to a number of related variables such as item difficulty, evaluator severity, and the challenge of the task. A third assessment instrument, referred to as the final examination, was administered at the end of the fourth week of instruction and is described more thoroughly below.

Design and procedure. This study was conducted as a part of a larger investigation, the GenScope Assessment Project, a National Science Foundation supported initiative that examined the effect assessment practices had on student engagement, motivation, and outcome performance (Hickey et al., 2002). A key aspect of the project was use of an exploratory computer software program, called GenScope™ (Horowitz & Christie, 2000) that supported the development of reasoning in genetics. An important attribute of the GenScope Assessment Project was the facilitation of assessment conversations (Duschl & Gitomer, 1997) as a way of determining what students understood about a subject based upon their interaction with the GenScope™ software and challenging near-transfer unit tests (Hickey et al., 2002).

The 2002 iteration of the GenScope Assessment Project was implemented in the classroom using teacher-directed and small group activities. These activities included use of paper and computer-based instructional materials intended to lead students through the GenScope curriculum. Typically, the teacher would introduce the day's curricular topic and model desired small-group activities before the entire class. In so doing, the teacher would often use an LCD projector panel to project a GenScope™ computer program image to illustrate operating principles before releasing the class to group activities that comprised the majority of the remaining instructional period. When released from

teacher-directed activities, students formed into stable membership, teacher-selected groups, and completed the day's activities, often using the GenScope™ computer program.

Students were participants who had returned parental permission slips which specifically permitted videotaping. In assigning participant groupings, in each of four different classrooms the teacher divided students into two groups, each with three members. These students were videotaped on four different occasions, yielding 32 videotapes (2 groups x 4 classrooms x 4 class sessions).

Across the period of the study, the teacher was present in all four classrooms during all class sessions and group membership remained unchanged. The teacher was an African-American male with an undergraduate life-science degree and was credentialed to teach secondary science. Additionally, for the purposes of this study, it was assumed that the effect of transactional discussion on outcome performance scores was the same across all individuals.

Students were administered a unit content test at the end of each of the four weeks of the GenScope curriculum. The fourth and last of these counted as a final examination. On the next class day following administration of each unit content test, students in their 3 member groups reviewed and graded their own performance.

In preparing students to review the first three unit content tests, the teacher modeled a "scripted" form of small-group review intended to foster formative assessment (Black, 1998). The teacher also gave examples of supporting assessment conversations which were seen as dialogue between group members that engaged them in many aspects of the ideas under consideration (Duschl & Gitomer, 1997).

During the first portion of each review period the teacher "walked" the class through the specific steps the students should use to evaluate their content tests. He also modeled desired group behavior and used researcher-provided answer explanations to scaffold student understanding of the unit content tests. This modeling of desired student interaction with course materials included instructions about how students were to use the answer explanations to determine their own grades.

Additionally, the teacher explained that it was not sufficient for students to simply note that they got a particular answer right or wrong as indicated on their answer sheets. Rather, students were instructed that they were to use their answer sheets and GenScope answer explanation sheets to scaffold a discussion in their groups about *why* they got an answer right or wrong. The GenScope answer explanation sheets were written at a higher conceptual level than the students' assumed reading level, but were provided to help explain the reasoning behind the questions in order to further scaffold student discussion.

The students were to work at their understanding of each problem on the test with the other members of the group until they each agreed that they understood the nature of their misunderstanding. The students were also instructed to take turns leading the discussion in order to assure that all group members were included in the conversation. The teacher then released the students to their respective groups to complete their evaluation of the remaining questions on each test using the answer and answer explanation sheets, and the group interaction methods he had modeled.

Unlike the first three unit content tests, the teacher and researchers graded the final examination. Otherwise, the teacher modeled the same methods used in reviewing unit content tests, and instructed the students to review the final examination using

researcher-provided answer explanation materials and the same small-group discussion techniques they used with the previous unit content tests.

Each of the student groups was videotaped as they reviewed and graded the results of their first three unit content tests, and as they reviewed the final examinations. Once begun, videotaping continued until the end of the period or until the teacher reconvened the class as a whole to review the day's activities.

Each of the videotaped participants wore separate microphones in order to allow the greatest possible signal separation as an aid to subsequent audio-visual analysis as well. Subjects were also asked to identify themselves as an aid to identification during audio-visual analysis.

Additionally, all participants were asked to ignore the video tape equipment as much as possible, and to participate in their group and class room activities in a normal manner. GenScope research staff members had little contact with group members during videotaping. However, in some instances the research staff did provide occasional instructions to group members in order to close the spacing of their chairs to improve visual coverage of their chairs during videotaping. In other instances the research staff answered specific curriculum related questions.

Coding and Scoring

Thirty-two videotapes of GenScope feedback activity were reviewed for the purposes of this study, with each videotape lasting approximately the entire class session, about 50 minutes. The primary researcher and a graduate student associated with the GenScope Assessment Project each transcribed 16 of the videotapes.

Transcription of each videotape was made in a script format and began once the videotape showed the classroom teacher releasing the class for group activities. Transcription did not begin until there were no adults participating in the videotaped group's activities, but included adult participation if they entered the scene after videotaping began. Only the first ten minutes of each group's activities were transcribed since a review of videotapes suggested that the most informative conversation occurred during the early portion of group activities.

The research centered on student participation in transactive discussion (Berkowitiz & Gibbs, 1983). The unit of analysis was the conversational turn which Kruger (1993) defined as each time a subject spoke without interruption. Each conversational turn was coded as reflecting either a transact or other form of communications (*non-transact*), with all utterances being coded as transactive or non-transactive. If it was not a transact, it was not further coded.

Following Kruger and Tomasello (1986), if a conversational turn was coded as a transact, it was further coded by type of transact and the orientation towards one's own ideas or others' ideas. As a result, each of the examples of transactive discussion was placed into one of the following categories: question, self-oriented; question, other-oriented; statement, self-oriented; statement, other-oriented; response, self-oriented; and response, other-oriented.

Observer reliability. The primary researcher transcribed and coded all group activities for transactive and non-transactive discussion. Additionally, 20% of the transcripts, selected randomly, were independently coded to determine interrater reliability. As all data was categorical, Cohen's Kappa was calculated to insure against

chance agreement (Huck & Cormier, 1996). The original intent was that coders would strive to achieve an inter-rater reliability level of .75 which is seen as excellent by both Bakeman and Gottman (1997), and Fleiss (1981). However, after training on an agreed coding scheme, the primary researcher and independent coder were only able to achieve an interrater reliability level of .61, a good, but not excellent level of agreement (Fleiss, 1981).

The initial interrater reliability was an initial point of concern. However, it was observed, in line with Bakeman and Gottman (1997), that the large number of non-transactive discussion utterances "undervalued" the relatively few transactive discussion utterances (55 of 1095 utterances in the sampled transcripts) in their various forms. This reduced what should have been a 7-category scheme (Question, Self-Oriented; Question, Other-Oriented; Statement, Self-Oriented; Statement, Other-Oriented; Response, Self-Oriented; Response, Other-Oriented; and non-transactive utterance) to a 2-category scheme (transactive utterance and non-transactive utterance). This also resulted in a skewed distribution that produced low values of kappa and denied credit for the coders' appropriate recognition that transactional statements did not occur when non-transactional dialogue did occur (probability of observation = .95).

A second Cohen's Kappa was computed mid-way through the coding process in order to check for drift between the primary coder and the independent coder. The new Cohen's Kappa was .62, once again less than the sought for .75 level of agreement, but was seen as good, given that, as before, the large number of non-transactive discussion utterances "undervalued" the relatively few transactive discussion utterances (60 of 1224 utterances) and resulted in low values of kappa.

Table 2 recapitulates the research questions. Table 2 also lists the type of analysis used, and the variables used in this study.

Table 2

Research Questions

Research Question	Type of Analysis	Variables Used
Research Question 1: What is the amount, type, and orientation of transacts that students generate overall during their feedback discussions?	Descriptive analysis	Self-Oriented Question Transacts Other-Oriented Question Transacts Self-Oriented Response Transacts Other-Oriented Response Transacts Self-Oriented Statement Transacts Other-Oriented Statement Transacts Transactive Utterances Non-Transactive Utterances Transactive Utterances Total Utterances
Research Question 2: Do students improve from pretest to posttest on scores of the SAT and NewWorm assessments?	Repeated measures	NewWorm Pretest SAT Pretest NewWorm Posttest SAT Posttest
Research Question 3: Does the total number of transacts increase over time?	Repeated measures	Transactive Utterances
Research Question 4: Will other-oriented transacts and question transacts become more frequent across the four feedback assessments examined in the GenScope research?	Repeated measures	Other-Oriented Transacts Question Transacts
Research Question 5: Is there a relationship between the number of other-oriented transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?	Correlation	Other-Oriented Transacts Final Examination NewWorm Pretest SAT Pretest NewWorm Posttest SAT Posttest
Research Question 6: Is there a relationship between the number of question transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?	Correlation	Question Transacts NewWorm Pretest SAT Pretest NewWorm Posttest SAT Posttest
Research Question 7: Is there a relationship between the total number of transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?	Correlation	Transactive Utterances NewWorm Pretest SAT Pretest NewWorm Posttest SAT Posttest

CHAPTER 4

RESULTS

Analysis Plan

The strategy of this analysis was to examine the impact of transacts uttered across the feedback sessions on the final examination, and NewWorm and SAT posttest assessments. However, the final examination was administered before feedback session 4 occurred. As a result, feedback session 4 had no effect on the final examination. For the analysis presented here, when examining the effects of transacts on the final examination, data from feedback sessions 1-3 were used. When examining the effects of transacts on the NewWorm and SAT posttest assessments, data from feedback sessions 1-4 were used.

Analysis of the Research Questions

Research Question 1

What is the amount, type, and orientation of transacts that students generate overall during their feedback discussions?

A descriptive analysis was conducted that focused on describing and graphing the incidence of all of the transactive discussion utterances made by the 24 students observed in this study. A count was made of the number and type of transacts made by each videotaped student.

The number of transactive utterances produced across feedback sessions 1 through 4 is presented in Table 3. Data are presented for utterances made for each

feedback session. In the last two columns of Table 3, the summary for feedback sessions 1-4 presents the number of transacts that were made by participants prior to the administration of the NewWorm and SAT posttest assessment outcome measures. The summary for feedback sessions 1-3 presents the number of transacts that were made by participants prior to the administration of the final examination outcome measure.

Figure 1 provides a graph of the total number of conversational turns that were transactive during the four feedback sessions. This again shows the relative increase in the number of statement transacts, both self-oriented and other-oriented, that were made by participants during Feedback Session 2. As also can be seen, there was a decline in most forms of transacts during Session 3 and a continued decline of transacts during Session 4.

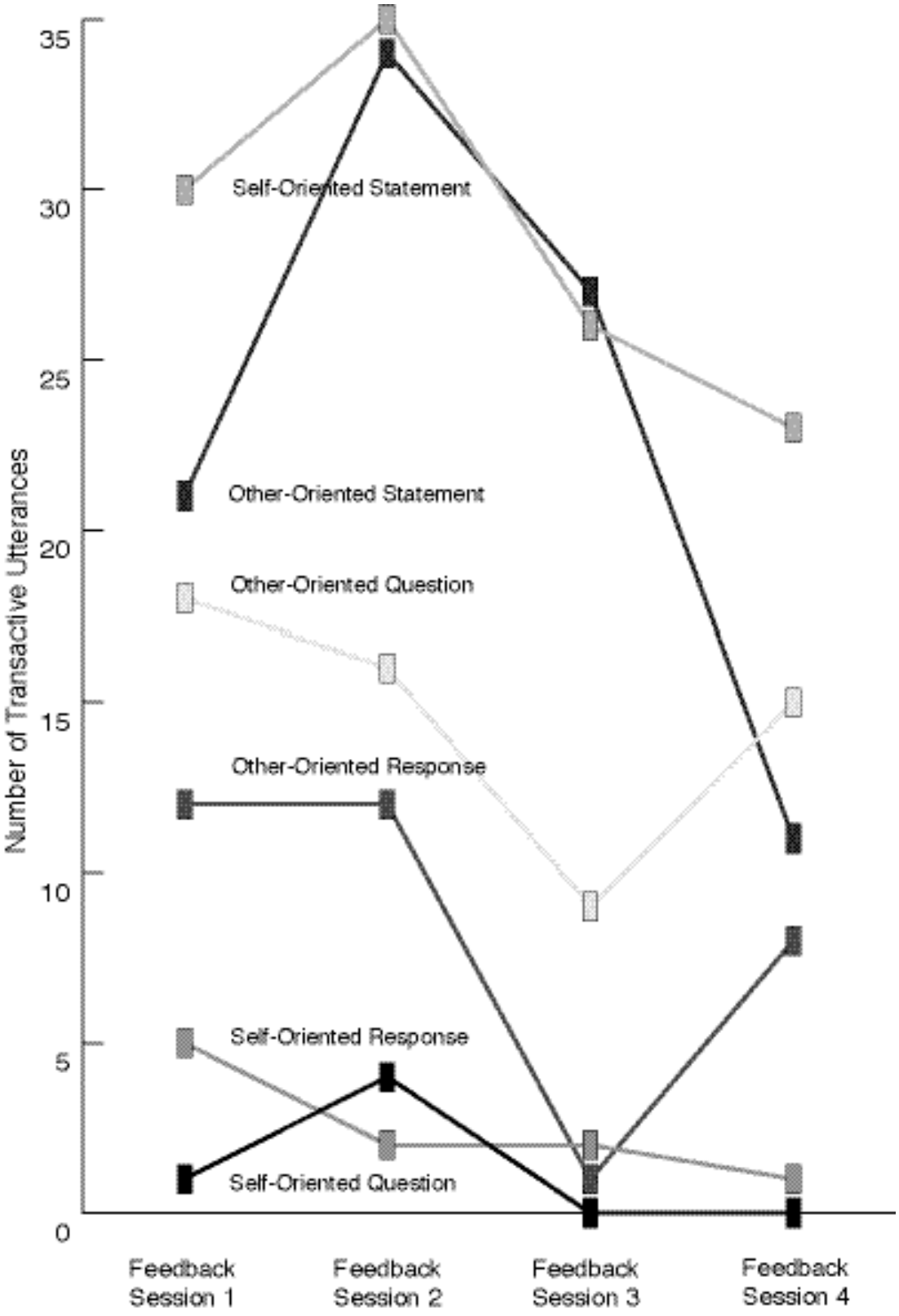
Table 3

Total Frequency of Conversational Turns That Are Transactive

Transactive utterances	Count by feedback session				Summary by feedback session	
	1	2	3	4	1 – 4	1 –3
Self-Oriented Question Transacts	1	4	0	0	5	5
Other-Oriented Question Transacts	18	16	9	15	58	43
Self-Oriented Response Transacts	5	2	2	1	10	9
Other-Oriented Response Transacts	12	12	1	8	33	25
Self-Oriented Statement Transacts	30	35	26	23	114	91
Other-Oriented Statement Transacts	21	34	27	11	93	82
Transactive Utterances	87	103	65	58	313	255
Non-transactive Utterances	814	1078	1027	1215	4134	2919
Total Utterances	901	1181	1092	1273	4447	3174

Note. Analyses of effects on NewWorm and SAT pretest-posttest comparisons are based on feedback sessions 1 through 4. Analyses of effects on final examination outcomes are based on feedback sessions 1 through 3.

Figure 1. Number of Conversational Turns That Are Transactive By Feedback Session



A different way of representing transacts is to report them as proportions of the total number of utterances. After comparison of results obtained in analyses involving incidence and proportions of transacts, both individually and when totaled, it was determined that for the purposes of this study, proportions did not offer any illustrative advantage for the data used in this investigation. Therefore, discussion of proportion has been dropped from further consideration in the study.

Research Question 2

Do students improve from pretest to posttest on scores of the SAT and NewWorm assessments?

Not all of the original 24 participants completed the NewWorm pretest and posttest assessments, and the SAT pretest and posttest assessments. Only those participants who completed both the NewWorm pretest and posttest assessments, and the SAT pretest and posttest assessments were evaluated in this research question. Table 4 reports the summary statistics for those participants whose NewWorm and SAT assessment scores were evaluated.

Table 4

Summary Statistics for NewWorm and SAT Pretest and Posttest Assessments

Variable	*Range*	n	*M*	SD
NewWorm Pretest	–3.20 to 58.26	20	38.08	14.57
NewWorm Posttest	42.25 to 72.33	20	54.28	7.26
SAT Pretest	25.21 to 60.09	20	42.02	13.31
SAT Posttest	25.21 to 67.88	20	52.99	11.97

A repeated measures ANOVA with 2 within subject levels (NewWorm Pretest and NewWorm Posttest) determined that the NewWorm Posttest was significantly different from the NewWorm Pretest, $F(1,19) = 36.43$, $p < .01$. Table 5 reports the repeated measure analysis of variance summary.

Table 5

Univariate Repeated Measure Analysis of Variance Summary for NewWorm Pretest and Posttest

Source	*DF*	*Sum of Squares*	*Mean Square*	*F*	*p*
Transacts	1	2623.59	2623.59	36.43	<.01
Error	19	1368.17	72.01		

A repeated measures ANOVA with 2 within subject levels (SAT Pretest and SAT Posttest) determined that the SAT Posttest was significantly different from the SAT Pretest, $F(1,19) = 9.51$, $p < .01$. Table 6 reports the repeated measure analysis of variance summary.

Table 6

Univariate Repeated Measure Analysis of Variance Summary for SAT Pretest and Posttest

Source	*DF*	*Sum of Squares*	*Mean Square*	*F*	*p*
Total Transacts	1	1204.62	1204.62	9.51	<.01
Error	19	2406.22	126.64		

Research Question 3

Does the total number of transacts increase over time?

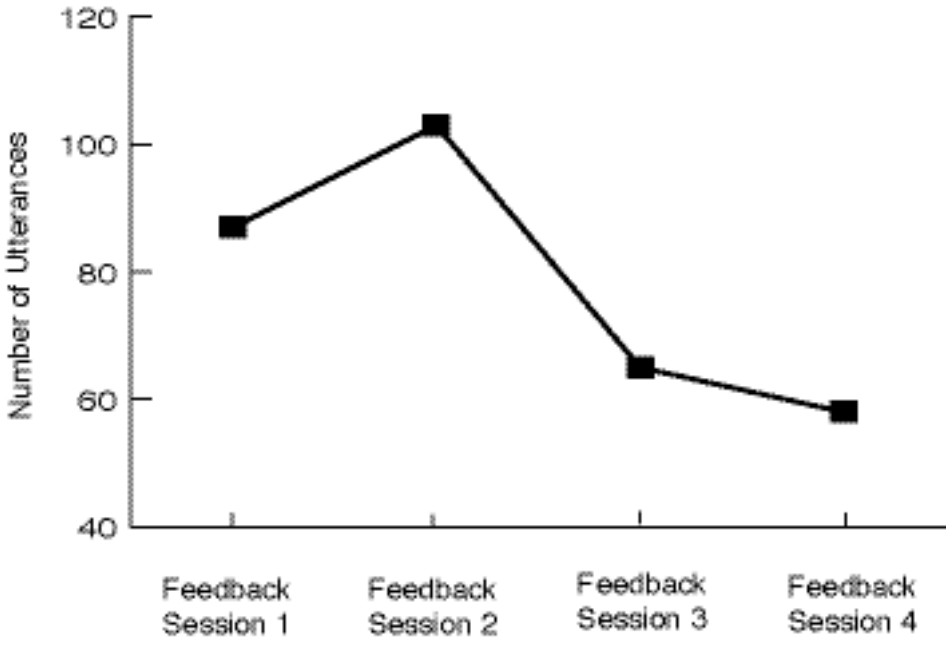
A repeated measures ANOVA with 4 within subject levels (feedback sessions 1, 2, 3, and 4) determined that there was a significant difference in the total number of transacts made between feedback sessions 1 through 4, $F(3,51) = 2.77$, $p = .05$. However, post hoc comparisons using Tukey's procedure (Stevens, 1990) failed to reveal a significant difference between the number of transacts made during feedback sessions 1 through 4. Table 7 reports the repeated measure analysis of variance summary. Figure 2 provides a graph of the total number of transacts across the four feedback sessions.

Table 7

Univariate Repeated Measure Analysis of Variance Summary for Total Number of Transacts

Source	*DF*	*Sum of Squares*	*Mean Square*	*F*	*p*
Total Transacts	3	56.82	18.94	2.77	.05
Error	51	348.93	6.84		

Figure 2. Graph of the Number of Transactive Utterances Across Feedback Sessions



Research Question 4

W*ill other-oriented transacts and question transacts become more frequent across the four feedback assessments examined in the GenScope research?*

Other-oriented transacts. A repeated measures ANOVA with 4 within subject levels (feedback sessions 1, 2, 3, and 4) determined that there was no statistical difference in the number of other-oriented transacts made between feedback sessions 1 through 4, $F(3,51) = 2.20$, $p = .10$, *NS*. Table 8 reports the repeated measure analysis of variance summary.

Table 8

Univariate Repeated Measure Analysis of Variance Summary for Other-Oriented Transacts

Source	*DF*	*Sum of Squares*	*Mean Square*	*F*	*p*
Other-Oriented Transacts	3	28.67	9.56	2.20	.10
Error	51	221.33	4.34		

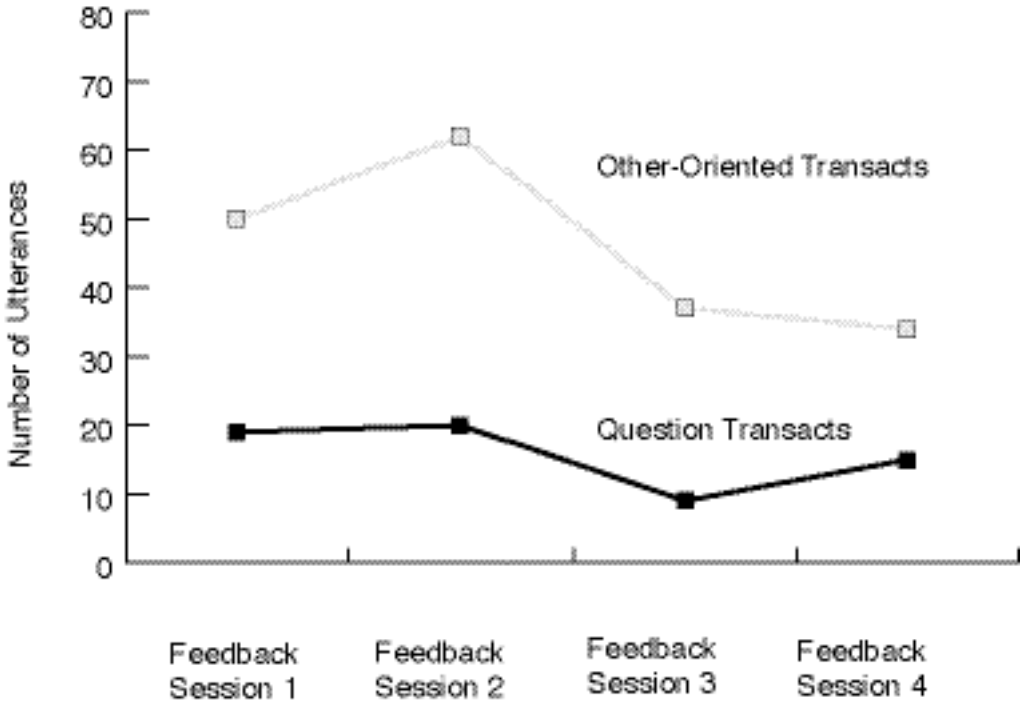
Question transacts. A repeated measures ANOVA with 4 within subject levels (feedback sessions 1, 2, 3, and 4) determined that there was no statistical difference in the number of question transacts made between feedback sessions 1 through 4, $F(3,51) = .74$, $p = .54$, *NS*. Table 9 reports the repeated measure analysis of variance summary. Figure 3 provides a combined graph of the total number of other-oriented transacts and question transacts seen across the four feedback sessions.

Table 9

Univariate Repeated Measure Analysis of Variance Summary for Question Transacts

Source	*DF*	*Sum of Squares*	*Mean Square*	*F*	*p*
Question Transacts	3	2.38	.79	.74	.54
Error	51	54.38	1.07		

Figure 3. Combined Graph of Other-Oriented and Question Transacts Seen Across Feedback Sessions 1 - 4



Research Question 5

Is there a relationship between the number of other-oriented transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?

This research question was analyzed using partial correlations in order to investigate the relationship between the number of other-oriented transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination. Effects of the NewWorm and SAT pretest performance were "partialed-out" in order to control for the effects of participant prior knowledge that may have effected performance on the posttest assessments and final examination.

As reported in Table 10, across feedback sessions 1 though 4, the relationship between the number of other-oriented transacts and performance on the NewWorm posttest was not significant when controlled for the NewWorm pretest, $r = .40$, $p = .09$, *NS*. Additionally, there does not appear to be a relationship between the number of other-oriented transacts and performance on the SAT posttest when controlled for the SAT pretest, $r = -.04$, $p = .87$, *NS*.

Across feedback session 1 through 3, the relationship between the number of other-oriented transacts and the final examination was not significant when controlled for the NewWorm pretest, $r = .43$, $p = .07$, *NS*. Additionally, the relationship between the number of other-oriented transacts and the final examination was not significant when controlled for the SAT pretest, $r = .44$, $p = .06$, *NS*.

Table 10

Partial Correlation Matrix of Other-Oriented Transacts Uttered Across All Feedback Sessions

Utterance	[a]*NewWorm*	[b]*SAT*	[c]*Final*	[d]*Final*
Other-Oriented	.40	-.04	.43	.44

Note. Analyses of effects on final examination outcomes are based on feedback sessions 1 through 3, NewWorm and SAT pretest-posttest comparisons are based on feedback sessions 1 through 4.
a – NewWorm posttest controlled for NewWorm pretest, b – SAT posttest controlled for SAT pretest, c – Final examination controlled for NewWorm pretest, d – Final examination controlled for SAT pretest.

Research Question 6

Is there a relationship between the number of question transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?

In a manner similar to research question 5, this research question was analyzed using partial correlations. In this case, the partial correlations were used to investigate the relationship between the number of question transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination.

As reported in Table 11, across feedback sessions 1 though 4, the relationship between the number of question transacts and performance on the NewWorm posttest was not significant when controlled for the NewWorm pretest, $r = .43$, $p = .07$, *NS*. Additionally, there does not appear to be a relationship between the number of question transacts and performance on the SAT posttest when controlled for the SAT pretest, $r = .00$, $p = .99$, *NS*.

Across feedback session 1 through 3, there does not appear to be a relationship between the number of question transacts and the final examination when controlled for the NewWorm pretest, $r = .13$, $p = .59$, *NS*. Nor does there appear to be a relationship between the number of question transacts and the final examination when controlled for the SAT pretest, $r = .16$, $p = .51$, *NS*.

Table 11

Partial Correlation Matrix of Question Transacts Uttered Across All Feedback Sessions

Utterance	[a]*NewWorm*	[b]*SAT*	[c]*Final*	[d]*Final*
Questions	.43	.00	.13	.16

Note. Analyses of effects on final examination outcomes are based on feedback sessions 1 through 3, NewWorm and SAT pretest-posttest comparisons are based on feedback sessions 1 through 4.
a – NewWorm posttest controlled for NewWorm pretest, b – SAT posttest controlled for SAT pretest, c – Final examination controlled for NewWorm pretest, d – Final examination controlled for SAT pretest.

Research Question 7

Is there a relationship between the total number of transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?

In a manner similar to research question 5, this research question was analyzed using partial correlations. In this case, the partial correlations were used to investigate the relationship between the total number of transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination.

As reported in Table 12, across feedback sessions 1 though 4, there does appear to be a relationship between the total number of transacts and performance on the NewWorm assessment posttest, $r = .60$, $p < .01$. However, there does not appear to be a relationship between the total number of transacts and performance on the SAT posttest when controlled for the SAT pretest, $r = .05$, $p = .84$, *NS*.

Across feedback session 1 through 3, there does appear to be a significant relationship between the total number of transacts and the final examination when controlled for the NewWorm pretest, $r = .49$, $p = .03$. There also appears to be a significant relationship between the total number of transacts and the final examination when controlled for the SAT pretest, $r = .53$, $p = .02$.

Table 12

Partial Correlation Matrix of Transacts Uttered Across All Feedback Sessions

Utterance	[a]*NewWorm*	[b]*SAT*	[c]*Final*	[d]*Final*
Transacts	.60*	.05	.49*	.53*

Note. Analyses of effects on final examination outcomes are based on feedback sessions 1 through 3, NewWorm and SAT pretest-posttest comparisons are based on feedback sessions 1 through 4.
a – NewWorm posttest controlled for NewWorm pretest, b – SAT posttest controlled for SAT pretest, c – Final examination controlled for NewWorm pretest, d – Final examination controlled for SAT pretest.
*$\underline{p} < .05$. *All probability values are one-tailed.*

Table 13 recapitulates the results of the research questions and identifies the methods of analysis for each research question. Four thousand four hundred forty seven utterances were analyzed during the course of this investigation, of which 313 were transactive and 4134 were non-transactive. These utterances were further analyzed by type and orientation. A significant difference was found between participant NewWorm pretest and posttest assessment scores, and between participant SAT pretest and posttest assessment scores. No significant difference was found in the total number of transacts found across the feedback sessions. No significant differences were found in the number of other-oriented transacts and question transacts across feedback sessions. No significant difference was found in the relationship between the number of other-oriented transacts and performance on the NewWorm posttest assessment when controlled for the NewWorm pretest. No significant difference was found in the relationship between other-oriented transacts and performance on the SAT posttest assessment when controlled for the SAT pretest. No significant difference was found in the relationship between the number of other-oriented transacts and the final examination when controlled for the NewWorm pretest. No significant difference was found in the relationship between the

number of other-oriented transacts and the final examination when controlled for the SAT pretest. No significant difference was found in the relationship between the number of question transacts and performance on the NewWorm posttest when controlled for the NewWorm pretest. No significant difference was found in the relationship between other-oriented transacts and performance on the SAT posttest assessment when controlled for the SAT pretest. Additionally, no significant difference was found in the relationship between the number of question transacts and performance on the final examination when controlled for either the NewWorm pretest or the SAT pretests. A significant relationship was found between the total number of transacts and performance on the NewWorm posttest assessment when controlled for the NewWorm pretest. However, no significant difference was found in the relationship between other-oriented transacts and performance on the SAT posttest assessment when controlled for the SAT pretest. Additionally, a significant relationship was found between the total number of transacts and the final examination when controlled for either the NewWorm or SAT pretests

A number of additional analyses were undertaken outside the scope of the original seven research questions and are presented at Appendix A. These additional investigations explored the relationship between transactive discussion and outcome performance measures. Moreover, these additional investigations consisted of a repeated measures analysis of variance, partial correlations, multiple regressions, and a consideration of time on task effects.

Table 13

Summary of Research Questions

Research Question	Type of Analysis	Results
Research Question 1: What is the amount, type, and orientation of transacts that students generate overall during their feedback discussions?	Descriptive analysis	Participants made a total of 313 transactive utterances and 4134 non-transactive utterances during the course of the study. Across the study, they also uttered 5 self-oriented question transacts, 58 other-oriented question transacts, 10 self-oriented response transacts, 33 other-oriented response transacts, 114 self-oriented statement transacts, and 93 other-oriented statement transacts.
Research Question 2: Do students improve from pretest to posttest on scores of the SAT and NewWorm assessments?	Repeated measures	Participants improved significantly from pretest to posttest on both the NewWorm and SAT assessments.
Research Question 3: Does the total number of transacts increase over time?	Repeated measures	No significant difference was found over time in the total number of transacts.
Research Question 4: Will other-oriented transacts and question transacts become more frequent across the four feedback assessments examined in the GenScope research?	Repeated measures	No significant difference was found across the feedback sessions in the number of other-oriented transacts. No significant difference was found across the feedback sessions in the number of question transacts.
Research Question 5: Is there a relationship between the number of other-oriented transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?	Correlation	No significant difference was found in the relationship between the number of other-oriented transacts and performance on the NewWorm posttest when controlled for the NewWorm pretest. No significant difference was found in the relationship between the number of other-oriented transacts and performance on the SAT posttest. No significant difference was found in the relationship between the number of other-oriented transacts and the final examination when controlled for either the NewWorm pretest or the SAT pretest.
Research Question 6: Is there a relationship between the number of question transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?	Correlation	No significant difference was found in the relationship between the number of question transacts and performance on the NewWorm posttest when controlled for the NewWorm pretest. No significant difference was found in the relationship between the number of question transacts and performance on the SAT posttest. No significant relationship was found between the number of question transacts and performance on the final examination whether controlled for the NewWorm pretest or the SAT pretest.
Research Question 7: Is there a relationship between the total number of transacts and performance on the NewWorm and SAT posttest assessments, and performance on the final examination?	Correlation	A significant relationship was found between the total number of transacts and performance on the NewWorm posttest when controlled for the NewWorm pretest. However, no significant difference was found in the relationship between the total number of transacts and performance on the SAT post-test when controlled for the SAT pretests. A significant relationship was found between the total number of transacts and the final examination when controlled for either the NewWorm pretest or the SAT pretest.

CHAPTER 5

DISCUSSION

Overview

This study was intended to investigate the relationship between several forms of transactive discussion and performance on achievement measures related to an instructional program. In that process, the study answered seven research questions targeting the nature of the relationship between transactive discussion and participant performance. The first four research questions were descriptive in nature and sought to quantify pretest and posttest outcome performance, as well as the nature of discourse utterances. The last three research questions investigated the relationship between transactive discussion and performance on assessments of outcome performance.

The findings of this study suggested that there was a relationship between transactive utterances and performance on the outcome measures used in the study. However, the nature of that relationship was not that originally suggested in the initial review of the literature. That initial review indicated that other-oriented transactive discussion was associated with gains on outcome measures. Instead, the results of this study illustrated the relation of the total number of transacts in their positive affect on outcome performance.

General Discussion

The first research question described the types and orientation of transactive discussion observed during this study. Participants made 4447 utterances during the

course of the study. Of these, 4134 utterances were non-transactive. Only the remaining 313 utterances (7.04%) were transactive. Of the transactive utterances, 5 were self-oriented question transactive utterances (0.11%), 58 were other-oriented question transactive utterances (1.30%), 10 were self-oriented response transactive utterances (0.22%), 33 were other-oriented response transactive utterances (0.74%), 114 were self-oriented statement transactive utterances (2.56%), and 93 were other-oriented statement transactive utterances (2.09%).

Additionally, fewer types of all transactive utterances were made during the fourth and final feedback session than during the first feedback session. However, rather than a general downward trend in the number of all types of transactive utterances, three of the six forms of transactive utterances investigated here (self-oriented statement transacts, other-oriented statement transacts, and self-oriented question transacts) showed an increase in the number of transacts in the second feedback session compared to the first or following feedback sessions.

The investigation of research question 3 showed that contrary to expectations, there was a difference ($p = .05$) in the number of transacts across feedback sessions. However, a post hoc comparison using Tukey's procedure failed to reveal a significant difference between the number of transacts made during the feedback sessions.

This research question investigated the idea that participants would use more transacts over time. An increasing use of transacts over time was expected because it was believed that as participants became more familiar with study materials and methods, and learned to participate in a meaningful way (Hickey & McCaslin, 2000) in the activities of their peer and study groups, their discourse would become more transactive. A similar

perspective was noted by Moreland et al. (1996) in their observation that transactive behavior was likely to increase as groups worked together. Similarly, Hakkarainen (2003) observed that explanatory activities increased with longer participation in a classroom environment. That the number of transactive utterances did not continually increase across the remaining feedback sessions was unexpected and is perhaps only partially explained post-hoc by consideration of the nature and difficulty of study materials seen in the different feedback sessions.

Additionally, the final examination was sequenced between the third feedback session, during which the participants reviewed their performance on the assessment judged as being the most difficult, and the fourth feedback session. As the participants appeared to be aware that they were simply reviewing their final examinations during the fourth feedback session and that it had already been graded by the teacher and GenScope Assessment Project experimenters, they may not have been as motivated to participate in group activities. As a result, although the number of transactive utterances increased during the second feedback session, overall there was a declining trend in the number of transactive utterances across the feedback sessions.

The investigation of the fourth research question showed that there was no significant difference in the number of other-oriented transacts and the number of question transacts made across the four feedback sessions. However, as researchers (Kruger, 1993; Azmitia & Montgomery, 1993) had observed that other-oriented transacts contributed to cognitive gain and problem solution, similar performance gains were expected. Moreover, it was expected that since gains in outcome performance had been demonstrated in the NewWorm and SAT posttests, that other-oriented transacts would be

found to increase across the feedback sessions as the participants became more familiar with study materials and methods as discussed previously.

Additionally, Scardamalia and Bereiter (1992), and Hakkarainen (2003) found that as children became more experienced in an instructional environment, they asked questions intended to help them understand that environment. As a result of these studies, it was expected that the number of participant questions would be found to increase across the feedback sessions as they continued to extend their knowledge. That no significant difference was found was unexpected.

Correlational analyses were conducted to answer the final three research questions. The first of these addressed the fifth research question and indicated that the relationship between other-oriented transacts, and the NewWorm posttest and the final examination was not significant. However, it had been expected that there would be a significant positive relationship between other-oriented transacts and outcome performance because of research such as Azmitia and Montgomery's (1993) observation of a positive link between other-oriented transacts and cognitive growth.

The second set of correlations addressed the sixth research question and indicated that the relationship between the number of question transacts, and the NewWorm posttest was not significant. This result was unexpected, as engaged discussion such as that represented by the use of question transacts was seen as being predictive of success in outcome performance (Kruger, 1992, 1993) and was therefore anticipated. This is also contrary to the observations of Scardamalia and Bereiter (1992), and Hakkarainen (2003) that children asked more questions as they became more experienced in an environment.

Finally, a third set of correlations addressed the seventh research question and indicated that there does appear to be a significant positive relationship between the total number of transactive utterances and performance on the NewWorm assessment posttest. Additionally, partial correlations reflecting transactive utterances across the first three feedback sessions showed a significant positive relationship with the final examination when controlled for the NewWorm pretest or SAT pretest.

Some researchers suggest there is relatively little support for the idea that peer interaction facilitates improvements in cognitive performance (Forman & Cazden, 1994). However, in contrast, Slavin (1996) supported the idea that discourse and other aspects of collaborative activities contributed to participant achievement. Similarly, Damon (1984), in contrasting Piagetian and Vygotskian perspectives on peer-based education, argued that peer discourse encourages cognitive development. Additionally, Mugny and Doise (1978) specifically noted that group performance exceeds individual performance and is similar to Lonning's (1993) observation that students in a 10th grade general science classroom showed gains on outcome performance as a result of peer collaboration. More recently, other researchers as well have noted the contribution of social interaction in a collaborative learning environment to positive outcome performance (Arvaja, Häkkinen, Pasku-Puttonen, & Eteläpelto, 2002, van Boxtel, van der Linden, & Kanselaar, 2000a; Vaughn, 2002).

Limitations of the Study

Several limitations may constrain the generalization of the study. As the research presented here is part of the larger body of research associated with the GenScope Assessment Project, it should generalize to the common population in the same manner in the same manner as that larger project. However, unlike the full population of the

GenScope Assessment Project, all participants in this study were minority students. This suggests it might not be possible to generalize the specific findings of the study to the larger population as minority groups have shown to demonstrate different discourse practices (Michaels, 1981) and achievement orientations (Steinberg, Dornbusch, & Brown, 1992).

Another limitation to this study is the lack of a significant sample size (Osborne, 2000), which limits its generalizability. Given a larger sample size, more powerful statistical methodologies such as latent change analysis (see for example Raykov and Marcoulides, 2000) might be employed to more fully examine the relation of transactional discussion to the outcome performance observed here.

A further limitation to the study is that the process of using answer explanation sheets and other GenScope Assessment Project materials may have reduced the incidence and effectiveness of transactive discussion observed in the course of the study. As noted by van Boxtel, van der Linden, & Kanselaar (2000) in their examination of the use of textbooks in physics instruction, such materials may moderate levels of discourse. This concern is appropriate in that a large proportion of transactive discussion research focuses on moral reasoning that did not appear to involve the direct use of supporting texts and textbooks (see especially Berkowitz, 1980a, 1980b, 1985; Berkowitz & Gibbs, 1983; Berkowitz et al., 1980, 1987; Damon & Killen, 1982; Pratt, Arnold, & Diessner, 1999).

Implications for Education

This study has shown that transactive discussion is present in the classroom and in that regard, it is similar to other studies that also observed the presence of transactive discourse in the classroom (see for example Berkowitz and Simmons, 2003). Although it

remains to be demonstrated that it is a property of all classrooms, as suggested by Wegner (1985), transactive systems may well be an emergent property (Johnson, 2001) of all groups that work closely together.

Additionally, the results of this study have shown that there is a significant relationship between the amount of transactive discourse and performance on outcome measures. The major implication of this finding is that classroom teachers should pay attention to the types of discourse occurring in their classrooms as the presence, absence, or quantity of transactive discourse is related to the outcome performance of students in that class. Moreover, in the context of the larger GenScope Assessment Project, teachers should seek ways to foster development of substantive discourse around academic subjects in the classroom.

This study also suggests that peer collaboration offers a unique opportunity to harness the social processes implied in transactive discussion, and peer collaboration for instructional purposes. Additionally, the use of such methods may offer advantages in minority student classrooms as suggested in research (Reid, 1992; Vaughn, 2002) that showed minority participants benefited from peer activities in mathematics.

Considering future research proceeding from this study, it should be replicated with a larger sample size and different demographics with a view towards establishing the generalizability of the results.

REFERENCES

Acton, G. S. (2003). What is good about Rasch measurement. *Rasch Measurement Transactions, 16*, 902-903.

Anton, M. (1999). The discourse of a learner-centered classroom: Sociocultural perspectives on teacher-learner interaction in the second-language classroom. *The Modern Language Journal, 83*(iii), 303-318.

Arvaja, M., Häkkinen, P., Pasku-Puttonen, H., & Eteläpelto, A. (2002). Social processes and knowledge building during small group interaction in a school science project. *Scandinavian Journal of Educational Research, 46*(2), 161-179.

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific meaning. *Social Development, 2*(3), 202-221.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction* (2nd ed.). New York: Cambridge University Press.

Ball, D. L. (1996). Teacher learning and the mathematics reforms. *Phi Delta Kappan, 77*(7), 500-508.

Berkowitz, M. W. (1980a). The role of transactive discussion in moral development: The history of a six-year program of research-part I. *Moral Education Forum, 5*(3), 13-26.

Berkowitz, M. W. (1980b). The role of transactive discussion in moral development: The history of a six-year program of research-part II. *Moral Education Forum, 5*(4), 15-27.

Berkowitz, M. W. (1985). The role of discussion in moral education. In M. W. Berkowitz & F. Oser (Eds.), *Moral education: Theory and application* (pp. 197-218). Hillsdale, NJ.

Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly, 29*(4), 399-410.

Berkowitz, M. W., & Gibbs, J. C. (1985). The process of moral conflict resolution and moral development. *New Directions for Child Development, 29*, 71-84.

Berkowitz, M. W., & Simmons, P. (2003). Integrating science education and character education: The role of peer discussion. In D. L. Zeidler (Ed.), *The role of moral reasoning on socioscientific Issues and Discourse in Science Education* (pp. 117-138). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Berkowitz, M. W., Gibbs, J. C., & Broughton, J. M. (1980). The relation of moral judgment stage disparity of developmental effects of peer dialogues. *Merrill-Palmer Quarterly, 26*(4), 341-357.

Berkowitz, M. W., Oser, F., & Althof, W. (1987). The development of sociomoral discourse. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Moral development through social interaction* (pp. 322-352). New York: Wiley.

Black, P. J. (1998). *Testing: Friend or foe? The theory and practice of assessment and testing*. London: The Falmer Press.

Black, P. J., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-75.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palinscar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting, the learning. *Educational Psychologist, 26*(3&4), 369-398.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141-178.

Brown, A. L. (1994). The advancement of learning. *Educational Researcher, 23*(8), 4-12.

Cameron, L. (2002). Metaphors in the learning of science: A discourse focus. *British Educational Research Journal, 28*(5), 673-688.

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

Cazden, C. B., & Beck, S. W. (2003). Classroom discourse. In A. C. Graesser, M. A. Gernsbacher & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 165-197). Mahwah, NJ: Lawrence Erlbaum Associates.

Chandler, P., & Sweller, J. (1989). *Cognitive load theory and the format of instruction*. Sydney, Australia: University of New South Wales.

Chinn, C. A., & Anderson, R. C. (1998). The structure of discussions that promote reasoning. *Teachers College Record, 100*(2), 315-368.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Cook, J. (2001a). The role of dialogue in computer-based learning on observing learning: An evolutionary approach to theory. Retrieved 20 December, 2001, from http://www-jime.open.ac.uk/2001/cook/cook-t.html.

Cook, K. (2001b). *Educational time factors*. Portland, OR: Northwest Regional Educational Laboratory.

Crook, C. (1996). *Computers and the collaborative experience of learning*. London: Routledge.

Damon, W. (1984). Peer education: The untapped potential. *Journal of Applied Developmental Psychology, 5*, 331-343.

Damon, W., & Killen, M. (1982). Peer interaction and the process of change in children's moral reasoning. *Merrill-Palmer Quarterly, 28*(3), 347-367.

De Lisi, R., & Golbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 3-37). Mahwah, NJ: Lawrence Erlbaum Associates.

Ding, W. (2003). *A study of collaborative scientific discovery*. Unpublished Dissertation, Carnegie Mellon University, Pittsburg, PA.

Doise, W., & Mugny, G. (1979). Individual and collective conflicts of centrations in cognitive development. *European Journal of Psychology, 9*, 105-108.

Doise, W., Mugny, G., & Perret-Clermont, A. (1976). Social interaction and cognitive development: Further evidence. *European Journal of Social Psychology, 6*(2), 245-247.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher, 23*(7), 5-12.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 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.

Felton, M., & Kuhn, D. (2001). The development of argumentive discourse skill. *Discourse Processes, 32*(2&3), 135-153.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: John Wiley & Sons.

Forman, E. A. (1992). Discourse, intersubjectivity, and the development of peer collaboration: A Vygotskian approach. In L. T. Winegar & J. Valsinger (Eds.), *Children's development within social context* (Vol. 1, Metatheory and Theory, pp. 143-159). Hillsdale, NJ: Lawrence Erlbaum Associates.

Forman, E. A., & Cazden, C. B. (1994). Exploring Vygotskian perspectives in education: The cognitive value of peer interaction. In R. B. Ruddell, M. Rapp & H. Singer (Eds.), *Theoretical models and processes of reading*. Newark, DE: International Reading Association.

Forman, E. A., & Kraker, M. J. (1985). The social origins of logic: The contributions of Piaget and Vygotsky. *New Directions For Child Development, 29*, 23-39.

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction, 8*(6), 527-548.

Forman, E. A., & McPhail, J. (1993). Vygotskian perspective on children's collaborative problem-solving activities. In E. A. Forman & N. Minick (Eds.), *Contexts for learning: Sociocultural dynamics in children's development* (pp. 213-229). New York: Oxford University Press.

Gabriele, A. J., & Montecinos, C. (2001). Collaborating with a skilled peer: The influence of achievement goals and perceptions of partner's competence on the participation and learning of low-achieving students. *The Journal of Experimental Education, 69*(2), 152-178.

Ge, X., Yamashiro, K. A., & Lee, J. (2000). Pre-class planning to scaffold students for online collaborative learning activities. *Educational Technology & Society, 3*(3), Retrieved 18 February, 2003, from http://ifets.ieee.org/periodical/vol_2003_2000/b2002.html.

Gee, J. P. (1989). Literacy, discourse, and linguistics: Introduction. *Journal of Education, 171*(1), 5-18.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: MacMillan.

Hakkarainen, K. (2003). Emergence of progressive-inquiry culture in computer-supported collaborative learning. *Learning Environments Research, 6*(2), 199-220.

Hakkarainen, K., & Järvelä, S. (2002). Epistemology of inquiry and computer-supported collaborative learning. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2 carrying forward the conversation*. Mahwah, JNJ: Lawrence Erlbaum Associates.

Hambleton, R. K. (Ed.). (1985). *Item response theory*. Boston: Kluwer - Nijohoff Publishing.

Hartman, M. (1996). Thinking and learning in classroom discourse. *Volta Review, 98*(3). Retrieved January 20, 2005 from EBSCOHOST Research Database.

Herrenkohl, L. R., Palinscar, A. S., DeWater, L. S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *Journal of the Learning Sciences, 8*(3/4), 451-494.

Hickey, D. T. (1999). *Assessment, engagement, & epistemological reconciliation in a technology-supported genetics environment*. Atlanta, GA: Georgia State University.

Hickey, D. T., Kindfield, A. C. H., Horowitz, P., & Christie, M. A. (2003). *Integrating curriculum, instruction, assessment, and evaluation in a technology-supported genetics learning environment*. *American Educational Research Journal, 40*(2), 495-538.

Hickey, D. T., Kruger, A. C., Fredrick, L. D., Schafer, N. J., & Kindfield, A. C. H. (2002). *Balancing formative and summative science assessment practices: Year one of the GenScope assessment project.* Athens, GA: Learning & Performance Support Laboratory, University of Georgia.

Hickey, D. T., Kruger, A. C., Fredrick, L. D., Schafer, N. J., Zuiker, S. J., & Michael, M. A. (2004a). Design-based alignment of discourse, understanding, and achievement: The GenScope assessment project (pp. 1-73): University of Georgia.

Hickey, D. T., & McCaslin, M. (2000). Comparative views of context in engagement and learning. In S. Volet & S. Jarvela (Eds.), *Motivation in learning contexts: Theoretical and methodological implications*. Elmsford, NY: Pergamon Press.

Hickey, D. T., Zuiker, S. J., Taasoobshirazi, G., Schafer, N. J., & Michael, M. A. (2004). *Balancing formative and summative assessment to attain systemic validity: Three is the magic number*. Studies in Educational Evaluation, Manuscript submitted December, 2004.

Hicks, D. (1996). Discourse, learning, and teaching. In M. W. Apple (Ed.), *Review of research in education* (Vol. 21, pp. 49-95). Washington, DC: American Educational Research Association.

Hoel, T. L. (1997). Voices from the classroom. *Teaching and Teacher Education, 13*(1), 5-16.

Hogan, D. M., & Tudge, J. R. H. (1999). Implications of Vygotsky's theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 39-65). Mahwah, NJ: Lawrence Erlbaum Associates.

Horwitz, P., & Christie, M. A. (2000). Computer-based manipulatives for teaching scientific reasoning: An example. In M. J. Jacobson & R. B. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning* (pp. 163-191). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Howe, C., Tolmie, A., & Rodgers, C. (1990). Physics in the primary school: Peer interaction and the understanding of floating and sinking. *European Journal of Psychology of Education, 4*, 459-475.

Howe, C., Tolmie, A., & Rodgers, C. (1992). The acquisition of conceptual knowledge in science by primary school children: Group interaction and the understanding of motion down an incline. *British Journal of Developmental Psychology, 10,* 113-130.

Huck, S. W., & Cormier, W. H. (1996). *Reading statistics and research*. New York: HarperCollins Publishers, Inc.

Jimenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (1999). "Doing the lesson" or "doing the science": Argument in high school genetics. *Science Digest, 84*(6), 757-792.

Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York: Scribner.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.

King, A. (1995). Designing the instructional process to enhance critical thinking across the curriculum: Inquiring minds really do want to know: Using questioning to teach critical thinking. *Teaching of Psychology, 22*(1), 13-17.

King, A. (1997). Ask to Think-Tel why: A model of transactive peer tutoring for scaffolding higher level complex learning. *Educational Psychologist, 32*(4), 221-235.

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review, 10*(1), 57-74.

King, A., & Rosenshine, B. (1993). Effects of guided cooperative questioning on children's knowledge construction. *Journal of Experimental Education, 61*(2), 127-148.

Kruger, A. C. (1992). The effect of peer and adult-child transactive discussions on moral reasoning. *Merrill-Palmer Quarterly, 38*(2), 191-211.

Kruger, A. C. (1993). Peer collaboration: Conflict, cooperation, or both? *Social Development, 2*(3), 165-182.

Kruger, A. C., & Tomasello, M. (1986). Transactive discussion with peers and adults. *Developmental Psychology, 22*(5), 681-685.

Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review, 62*(2), 155-178.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning science thinking. *Science Education, 77*(3), 319-337.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*(4), 495-523.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction, 15*(3), 287-315.

Landsmann, L. T. (1991). The conceptualization of writing in the confluence of interactive models of development. In L. T. Landsmann (Ed.), *Culture, schooling, and psychological development* (pp. 87-111). Norwood, NJ: Ablex Publishing Corporation.

Lehtinen, E. (2003). Computer-supported collaborative learning: An approach to powerful learning environments. Retrieved July 20, 2004, from www.tml.hut.fi/Opinnot/T-110.556/2004/Materiaali/EditedLehtinenCSCL1.pdf

Leinonen, T., Virtanen, O., Hakkarainen, K., & Kligyte, G. (2002, January). *Collaborative discovering of key ideas in knowledge building*. Paper presented at the proceedings of Computer Support for Collaborative Learning (CSCL 2002) Conference, Boulder, CO.

Levine, J. M., & Resnick, L. B. (1993). Social foundations of cognition. *Annual Review of Psychology, 44*, 585-612.

Liang, W. L., Moreland, R. L., & Argote, L. (1995). Group versus individual training and group performance: The mediating factor of transactive memory. *Personality & Social Psychology Bulletin, 21*(4), 384-393.

Lonning, R. (1993). Effect of cooperative learning strategies on student verbal interactions and achievement during conceptual change instruction in 10th grade general science. *Journal of Research in Science Teaching, 30*(9), 1087-1101.

Messer, D. J., Joiner, R., Loveridge, N., Light, P., & Littleton, K. (1993). Influences on the effectiveness of peer interaction: Children's level of cognitive development and relative ability of partners. *Social Development, 2*(3), 279-294.

Michaels, S. (1981). "Sharing time": Children's narrative styles and differential access to literacy. *Language in Society, 10*, 423-442.

Miell, D., & MacDonald, R. (2000). Children's creative collaborations: The importance of friendship when working together on a musical composition. *Social Development, 9*(3), 348-369.

Miller, S. A., & Brownell, C. A. (1975). Peers, persuasion, and Piaget: Dyadic interaction between conservers and nonconservers. *Child Development, 46*, 992-997.

Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. L. Nye & A. M. Bower (Eds.), *What's social about social cognition?* (pp. 57-84). Thousand Oaks, CA: Sage Publications.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8*, 181-192.

Nathan, M. J., & Knuth, E. J. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21*(2), 175-207.

National Academy of Sciences. (1995a). National science education standards: An overview. In *National Science Education Standards*. Retrieved February 27, 2002, from http://books.Nap.Edu/html/nses/html/overview.html. Washington, DC: National Academy of Sciences.

National Academy of Sciences. (1995b). *National standards Signal New Era for Nation's K-12 Science Education: News From The National Research Council*. Retrieved March 16, 2002, from http://www.nap.edu/readingroom/books/nses/press.html

National Council of Teachers of Mathematics (1991). *Professional Standards for Teaching Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English, 25*(3), 261-290.

Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35*(2), 135-198.

Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science, 21*(2), 109-146.

Orsolini, M., & Pontecorvo, C. (1992). Children's talk in classroom discussions. *Cognition and Instruction, 9*(2), 113-136.

Osborne, J., Simon, S., & Erduran, S. (2002). Enhancing the quality of argumentation in school science, *Conference on Philosophical, Psychological, and Linguistic Foundations for Language and Science Literacy Research*. Victoria, BC: University of Victoria.

Osborne, J. W. (2000). Prediction in multiple regression. *Practical Assessment, Research & Evaluation, 7*(2). Retrieved February 12, 2004, from http://PAREonline.net/getvn.asp?v=2007&n=2002.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 49,* 345-375.

Perret-Clermont, A. (1980). *Social interaction and cognitive development in children* (C. Sherrad, Trans. Vol. 19). New York: European Association of Experimental Social Psychology.

Perret-Clermont, A., & Schubauer-Leoni, M. L. (1981). Conflict and cooperation as opportunities for learning. In P. Robinson (Ed.), *Communication in development* (pp. 203-233). London: Academic Press.

Phelps, E., & Damon, W. (1989). Problems solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology, 81*(4), 639-645.

Piaget, J. (1964). *The psychology of intelligence*. London: Routledge & Kegan Paul LTD.

Piaget, J. (1971). *Biology and knowledge; an essay on the relations between organic regulations and cognitive processes*. Chicago: University Press.

Piaget, J. (1995). *Sociological studies* (T. Brown, R. Campbell, N. Emler, M. Ferrari, M. Gribetz, R. Kitchner, W. Mays, A. Notari, C. Sheppard & L. Smith, Trans.). London: Routledge.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child* (H. Weaver, Trans.). New York: Basic Books, Inc.

Pontecorvo, C. (1993). Forms of discourse and shared thinking. *Cognition and Instruction, 11*(3 & 4), 189-196.

Pratt, M. W., Arnold, M. L., Pratt, A. T., & Diessner, R. (1999). Predicting adolescent moral reasoning from family climate: A longitudinal study. *Journal of Early Adolescence, 19*(2), 148-176.

Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Reid, J. (1992). The effects of cooperative learning with intergroup competition on the math achievement of seventh grade students. Retrieved January 6, 2004, from ERIC Document Reproduction Service No. ED355106

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences, 2*(3), 235-276.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research, 13*, 89-99.

Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge-based questioning by children. *Cognition and Instruction, 9*(3), 177-199.

Siegler, R. S., & Alibali, M. W. (2005). Piaget and the Neo-Piagetians. In *Children's thinking* (4 ed., pp. 78-117). Upper Saddle River, NJ: Prentice Hall.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*, 43-69.

Stein, M. K. (2001). Mathematical argumentation: Putting umph into classroom discussions. *Mathematics Teaching in the Middle School, 7*(2), 110-112.

Steinberg, L., Dornbusch, S. M., & Brown, B. B. (1992). Ethnic differences in adolescent achievement: An ecological perspective. *American Psychologist, 47*(6), 723-729.

Stevens, J. P. (1990). *Intermediate statistics*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Tao, P., & Gunstone, R. F. (1999). Conceptual change in science through collaborative learning at the computer. *International Journal of Science Education, 21*(1), 39-57.

Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology, 31*(2), 207-220.

Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Thorley, N. R., & Treagust, D. F. (1987). Conflict within dyadic interactions as a stimulant for conceptual change in physics. *International Journal of Science Education, 9*(2), 203-216.

Tolmie, A., Howe, C., Mackenzie, M., & Greer, K. (1993). Task design as an influence on dialogue and learning: Primary school group work with object flotation. *Social Development, 2*(3), 183-201.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16*, 495-552.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Philadelphia, PA: Open University Press.

Toulmin, S. E. (1958). *The uses of argument*. Cambridge, Great Britain: Cambridge at the University Press.

Tudge, J., & Winterhoff, P. (1993). Can young children benefit from collaborative problem solving? Tracing the effects of partner competence and feedback. *Social Development, 2*(3), 242-259.

van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000a). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*, 311-330.

van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000b). The use of textbooks as a tool during collaborative physics learning. *Journal of Experimental Education, 69*(1), 57-76.

Vaughn, W. (2002). Effects of cooperative learning on achievement and attitude among students of color. *The Journal of Educational Research, 95*(6), 359-364.

Vauras, M., IIskala, T., Kajamies, A., Kinnunen, R., & Lehtinen, E. (2003). Shared-regulation and motivation of collaborating peers: A case analysis. *Psychologia, 46*, 19-37.

Vermunt, J. D., & Vermetten, Y. J. (2004). Patterns in student learning: Relationships between learning strategies, conceptions of learning, and learning orientations. *Educational Psychology Review, 16*(4), 359-384.

Webb, N. M., & Palinscar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York: MacMillan.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York: Springer-Verlag.

Wegner, D. M., Giuliano, T., & Hertel, P. T. (1985). Cognitive interdependence in close relationships. In W. Ickes (Ed.), *Compatible and incompatible relationships*. New York: Springer-Verlag.

Wegner, D. M., Raymond, P., & Erber, R. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology, 61*(6), 923-929.

Williams, S. R., & Baxter, J. A. (1996). Dilemmas of discourse-oriented teaching in one middle school mathematics classroom. *The Elementary School Journal, 97*(1), 21-38.

APPENDIX A

Ancillary Analyses

Results

A series of additional analyses were undertaken beyond the original seven research questions in order to further investigate the relationship between transactive discussion and measures of outcome performance seen in the fifth, sixth, and seventh research questions. These additional analyses consisted of a repeated measures analysis of variance, partial correlations, multiple regressions, and a consideration of time on task effects.

Repeated Measures Analysis of Variance

In a visual inspection of Table 3 (*Total Frequency of Conversational Turns that are Transactive*) and Figure 1 (*Number of Conversational Turns That Are Transactive By Feedback Session*), it was noted that there appeared to be more self-oriented statement transacts and other-oriented statement transacts than other types of transacts. This suggested that *statement transacts* (the sum of self-oriented statement transacts and other-oriented statement transacts) might reflect a unitary measure, and that together there appeared to be more of them occurring during the second feedback session than in the other feedback sessions.

A repeated measures ANOVA with 4 within subject levels (feedback sessions 1, 2, 3, and 4) found no significant difference in the number of statement transacts made

between feedback sessions 1 through 4, $F(3,51) = 2.64$, $p = .06$, *NS*. Post hoc comparisons using Tukey's procedure found a significant difference in the number of statement transacts between feedback sessions 1 and 2, feedback sessions 1 and 4, feedback sessions 2 and 3, feedback sessions 2 and 4, and feedback sessions 3 and 4. However, no significant difference was found in the number of statement transacts between feedback sessions 1 and 3.

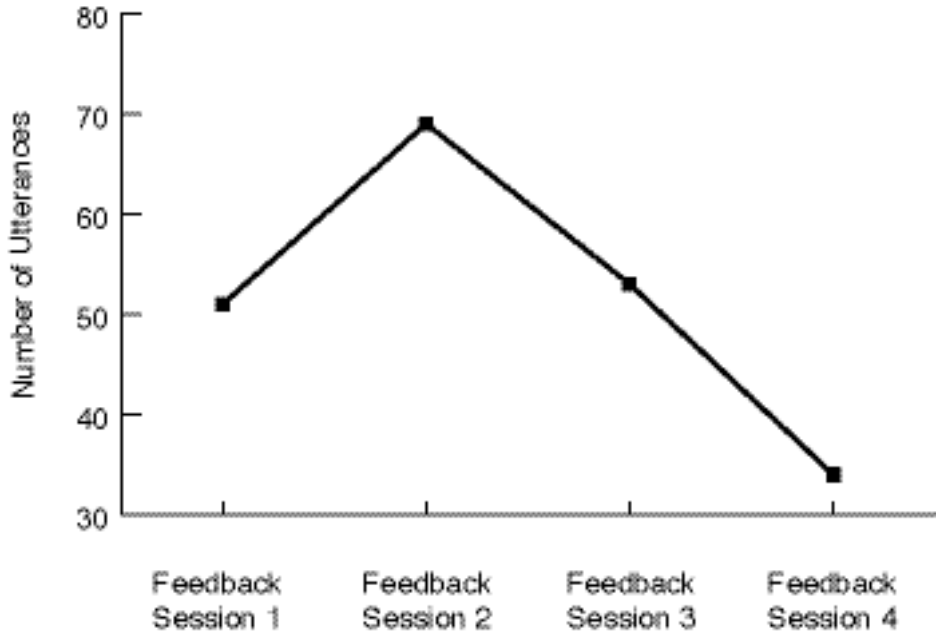
Table A1 reports the repeated measure analysis of variance summary. Figure A1 provides a graph of the total number of statement transacts across the four feedback sessions.

Table A1

Univariate Repeated Measure Analysis of Variance Summary for Statement Transacts

Source	*DF*	*Sum of Squares*	*Mean Square*	*F*	*p*
Statement Transacts	3	28.00	9.33	2.64	.06
Error	51	180.50	3.54		

Figure A1. Graph of the Number of Statement Transacts Across Feedback Sessions



Partial Correlations

In further examining the preceding repeated measures analysis of variance and Figure A1 (*Graph of the Number of Statement Transacts Across Feedback Sessions*) it was noted that the second feedback session had a significantly larger number of statement transacts when compared to the other feedback sessions. Partial correlations were used to investigate the relationship between statement transacts uttered during the second feedback session and performance on the NewWorm and SAT posttests, and performance on the final examination.

As reported in Table A2, across feedback session 2, no significant relationship was found between statement transacts and performance on the NewWorm assessment posttest, $r = .44$, $p = .07$, *NS*. No significant relationship was found between statement transacts and performance on the SAT posttest when controlled for the SAT pretest, $r = .18$, $p = .46$, *NS*. However, there appears to be a significant relationship between

statement transacts and performance on the final examination when controlled for the NewWorm pretest, $r = .59$, $p < .01$. There also appears to be a significant relationship between statement transacts and performance on the final examination when controlled for the SAT pretest, $r = .67$, $p < .01$.

Table A2

Partial Correlation Matrix of Statement Transacts Uttered During Feedback Session 2

Utterance	[a]*NewWorm*	[b]*SAT*	[c]*Final*	[d]*Final*
Transacts	.44	.18	.59*	.67*

Note. Analyses of effects on final examination outcomes are based on feedback sessions 1 through 3, NewWorm and SAT pretest-posttest comparisons are based on feedback sessions 1 through 4.
a – NewWorm posttest controlled for NewWorm pretest, b – SAT posttest controlled for SAT pretest, c – Final examination controlled for NewWorm pretest, d – Final examination controlled for SAT pretest.
* $p < .01$. All probability values are two-tailed.

Hierarchical Multiple Regressions

Based on the relationships suggested in the Research Questions 5, 6, 7, and their correlation matrices, two series of hierarchical multiple regression equations were modeled to further describe the effect of transactive utterances on outcome performance measures. In all of the hierarchical multiple regression equations modeled for these studies, the NewWorm and SAT pretest predictors were individually entered before transactive utterance predictors.

Hierarchical multiple regression equations describing feedback session 2. The first series of hierarchical multiple regression equations examined relationships between transacts, the NewWorm posttest, and the final examination seen during the second feedback session. These relationships were examined in an effort to further explain the effect of statement transacts seen during the second feedback session which were

explored in the preceding ancillary analyses repeated measures and partial correlation investigations. In this series of hierarchical multiple regression equations, eight equations were found to be significant when modeled to describe the effects of total transacts and statement transacts uttered during feedback session 2 on outcome measures.

As shown in Table A3, two of eight equations described the effect of the total number of transacts on the NewWorm posttest when controlled for the NewWorm pretest. The first equation created was significant ($F(2,18) = 6.97$, $p = .01$) and accounted for 47% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .12$, *NS*) and initially accounted for 33% of the variance. The total number of transacts uttered during feedback session 2 used as a predictor was not significant ($p = .06$, *NS*) and accounted for an additional 14% of the variance.

The second equation created was significant ($F(2,18) = 6.69$, $p = .01$) and accounted for 46% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .06$, *NS*) and accounted for 33% of the variance. The total number of statement transacts uttered during feedback session 2 used as a predictor was not significant ($p = .07$, *NS*) and accounted for an additional 13% of the variance.

Table A3

Hierarchical Multiple Regression Analyses Across Feedback Session 2 – NewWorm Posttest

		Model					Predictor's Contribution	
Outcome Measure	Predictor	F	*p*	*df*	R^2	$R^2\Delta$	*t*	*p*
[a]NewWorm	$NewWorm_{pretest}$	6.97	.01	2,18	.47	.33	1.63	.12
	Total Transacts					.14	2.05	.06
	$NewWorm_{pretest}$	6.69	.01	2,18	.46	.33	2.02	.06
	Statement Transacts					.13	1.96	.07

Note. a – NewWorm posttest controlled for NewWorm pretest

As shown in Table A4, three of eight equations described the effect of the total number of transacts on the final examination when controlled for the NewWorm pretest. The first equation created was significant ($F(2,20) = 4.85$, $p = .02$) and accounted for 35% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .54$, *NS*) and accounted for 14% of the variance. The total number of transacts uttered during feedback session 2 used as a predictor was significant ($p = .03$) and accounted for an additional 21% of the variance.

The second equation created was significant ($F(2,18) = 6.83$, $p < .01$) and accounted for 43% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .37$, *NS*) and accounted for 14% of the variance. The total number of

statement transacts uttered during feedback session 2 used as a predictor was significant (p = .01) and accounted for an additional 30% of the variance.

The third equation created was significant (F(2,20) = 4.94, p = .02) and accounted for 35% of the variance. The NewWorm pretest score used as a predictor was not significant (p = .31, *NS*) and accounted for 14% of the variance. The total number of self-oriented statement transacts uttered during feedback session 2 used as a predictor was significant (p = .02) and accounted for an additional 22% of the variance.

Table A4

Hierarchical Multiple Regression Analyses Across Feedback Session 2 – Final Examination Controlled for NewWorm Pretest

		Model					Predictor's Contribution	
Outcome Measure	Predictor	F	*p*	*df*	R^2	$R^2\Delta$	*t*	*p*
[a]Final	$NewWorm_{pretest}$	4.85	.02	2,20	.35	.14	.63	.54
	Total Transacts					.21	2.44	.03
	$NewWorm_{pretest}$	6.83	.01	2,18	.43	.14	.92	.37
	Statement Transacts					.30	3.06	.01
	$NewWorm_{pretest}$	4.94	.02	2,20	.35	.14	1.05	.31
	Self-Oriented Statement Transacts					.22	2.47	.02

Note. a – Final examination controlled for NewWorm pretest

As shown in Table A5, three of eight equations described the effect of the total number of transacts on the final examination when controlled for the SAT pretest. The first equation created was significant ($F(2,20) = 4.91$, $p = .02$) and accounted for 35% of the variance. The SAT pretest used as a predictor was not significant ($p = .50$, *NS*) and accounted for 8% of the variance. The total number of statement transacts uttered during feedback session 2 was significant ($p = .01$) and accounted for an additional 28% of the variance.

The second equation created was significant, ($F(2,18) = 8.60$, $p < .01$) and accounted for 49% of the variance. The SAT pretest used as a predictor was not significant ($p = .10$, *NS*) and accounted for 8% of the variance. The total number of statement transacts used as a predictor was significant ($p < .01$) and accounted for 41 % of the variance.

The third equation created was significant, ($F(2,20) = 8.88$, $p < .01$) and accounted for 50% of the variance. The SAT pretest used as a predictor was significant ($p = .02$) and accounted for 8% of the variance. The total number of self-oriented statement transacts uttered during feedback session 2 used as a predictor was significant ($p < .01$) and accounted for an additional 42% of the variance.

Table A5

Hierarchical Multiple Regression Analyses Across Feedback Session 2 – Final Examination Controlled for SAT Pretest

		Model					Predictor's Contribution	
Outcome Measure	Predictor	F	*p*	*df*	R^2	$R^2\Delta$	*t*	*p*
[a]Final	$SAT_{pretest}$	4.91	.02	2,20	.35	.08	.69	.50
	Total Transacts					.28	2.77	.01
	$SAT_{pretest}$	8.60	<.00	2,18	.49	.08	1.72	.10
	Statement Transacts					.41	3.80	<.00
	$SAT_{pretest}$	8.88	<.00	2,20	.50	.08	2.55	.02
	Self-Oriented Statement Transacts					.42	3.87	<.00

Note. a – Final examination controlled for SAT pretest.

Hierarchical multiple regression equations across all feedback sessions. The second series of hierarchical multiple regression equations examined the relationships between transacts, the NewWorm posttest, and the final examination seen during all feedback sessions. In this series of hierarchical multiple regression equations, ten equations were found to be significant when modeled to describe the effect of the total number of transacts uttered across all feedback sessions on outcome measures.

As shown in Table A6, two of ten equations described the effect of the total number of transacts uttered on the NewWorm posttest when controlled for the NewWorm

pretest. The first equation created was significant ($F(2,19) = 15.76$, $p < .01$) and accounted for 65% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .13$, *NS*) and accounted for 33% of the variance. The total number of self-oriented transacts uttered during feedback session 1 through 4 was significant ($p < .01$) and accounted for an additional 32% of the variance.

The second equation created was significant ($F(2,19) = 14.12$, $p < .01$) and accounted for 62% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .10$, *NS*) and accounted for 33% of the variance. The total number of self-oriented statement transacts uttered during feedback session 1 through 4 was significant ($p < .01$) and accounted for 30% of the variance.

Table A6

Hierarchical Multiple Regression Analyses Across All Feedback Sessions – NewWorm Posttest

		Model					Predictor's Contribution	
Outcome Measure	Predictor	*F*	*p*	*df*	R^2	$R^2\Delta$	*t*	*p*
[a]NewWorm	$NewWorm_{pretest}$	15.76	<.01	2,19	.65	.33	1.57	.13
	Self-Oriented Transacts					.32	3.96	<.01
	$NewWorm_{pretest}$	14.12	<.01	2,19	.62	.33	1.73	.10
	Self-Oriented Statement Transacts					.30	3.67	<.01

Note. a – NewWorm posttest controlled for NewWorm pretest

As shown in Table A7, four of ten equations described the effect of the total number of transacts on the final examination when controlled for the NewWorm pretest. The first equation created was significant ($F(2,19) = 3.99$, $p = .04$) and accounted for 32% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .55$, *NS*) and accounted for 10% of the variance. The total number of transacts uttered during feedback session 1 through 3 used as a predictor was significant ($p = .03$) and accounted for an additional 22% of the variance.

The second equation created was significant ($F(2,19) = 4.41$, $p = .03$) and accounted for 34 % of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .54$, *NS*) and accounted for 10% of the variance. The total number of statement transacts uttered during feedback session 1 through 3 used as a predictor was significant ($p = .02$) and accounted for an additional 24% of the variance.

The third equation created was significant ($F(2,21) = 4.38$, $p = .03$) accounted for 32% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .39$, *NS*) and accounted for 16% of the variance. The total number of self-oriented statement transacts uttered during feedback session 1 through 3 used as a predictor was significant ($p = .05$) and accounted for 16% of the variance.

The fourth equation created was significant ($F(2,21) = 4.34$, $p = .03$) and accounted for 31% of the variance. The NewWorm pretest score used as a predictor was not significant ($p = .11$, *NS*) and accounted for 16% of the variance. The total number of other-oriented statement transacts uttered during feedback sessions 1 through 3 was significant ($p = .05$) and accounted for 15% of the variance.

Table A7

Hierarchical Multiple Regression Analyses Across All Feedback Sessions – Final Examination Controlled for NewWorm Pretest

		Model					Predictor's Contribution	
Outcome Measure	Predictor	*F*	*p*	*df*	R^2	$R^2\Delta$	*t*	*p*
[a]Final	$NewWorm_{pretest}$	3.99	.04	2,19	.32	.10	.61	.55
	Total Transacts					.22	2.32	.03
	$NewWorm_{pretest}$	4.41	.03	2,19	.34	.10	.62	.54
	Statement Transacts					.24	2.48	.02
	$NewWorm_{pretest}$	4.38	.03	2,21	.32	.16	.89	.39
	Self-Oriented Statement Transacts					.16	2.08	.05
	$NewWorm_{pretest}$	4.34	.03	2,21	.31	.16	1.68	.11
	Other-Oriented Statement Transacts					.15	2.06	.05

Note. a – Final examination controlled for NewWorm pretest

As shown in Table A8, four of ten equations described the effect of the total number of transacts on the final examination when controlled for the SAT pretest. The first equation created was significant ($F(2,19) = 3.98$, $p = .04$) and accounted for 32% of the variance. The SAT pretest score used as a predictor was not significant ($p = .56$, *NS*) and accounted for 6% of the variance. The total number of transacts uttered during

feedback session 1 through 3 was significant (p = .02) and accounted for an additional 26% percent of the variance.

The second equation created was significant (F(2,19) = 4.93, p = .02) and accounted for 37% percent of the variance. The SAT pretest score used as a predictor was not significant (p = .31, *NS*) and accounted for 6% of the variance. The total number of statement transacts seen across feedback sessions 1 through 3 used as a predictor was significant (p = .01) and accounted for an additional 31% percent of the variance.

The third equation created was significant (F(2,21) = 4.43, p = .03) and accounted for 32% percent of the variance. The SAT pretest score used as a predictor was not significant (p = .36, *NS*) and accounted for 8% percent of the variance. The total number of self-oriented statement transacts uttered during feedback sessions 1 through 3 was significant (p = .02) and accounted for 24% percent of the variance.

The fourth equation created by its two predictor variables was significant (F(2,21) = 4.03, p = .04) and accounted for 30% percent of the variance. The SAT pretest score used as a predictor was not significant (p = .14, *NS*) and accounted for 8% percent of the variance. The total number of other-oriented statement transacts uttered during feedback sessions 1 through 3 was significant (p = .02) and accounted for 22% of the variance.

Table A8

Hierarchical Multiple Regression Analyses Across All Feedback Sessions – Final Examination Controlled For SAT Pretest

		Model					Predictor's Contribution	
Outcome Measure	Predictor	F	*p*	*df*	R^2	$R^2\Delta$	*t*	*p*
[a]Final	$SAT_{pretest}$	3.98	.04	2,19	.32	.06	.60	.56
	Total Transacts					.26	2.56	.02
	$SAT_{pretest}$	4.93	.02	2,19	.37	.06	1.05	.31
	Statement Transacts					.31	2.89	.01
	$SAT_{pretest}$	4.43	.03	2,21	.32	.08	.93	.36
	Self-Oriented Statement Transacts					.24	2.60	.02
	$SAT_{pretest}$	4.03	.04	2,21	.30	.08	1.53	.14
	Other-Oriented Statement Transacts					.22	2.46	.02

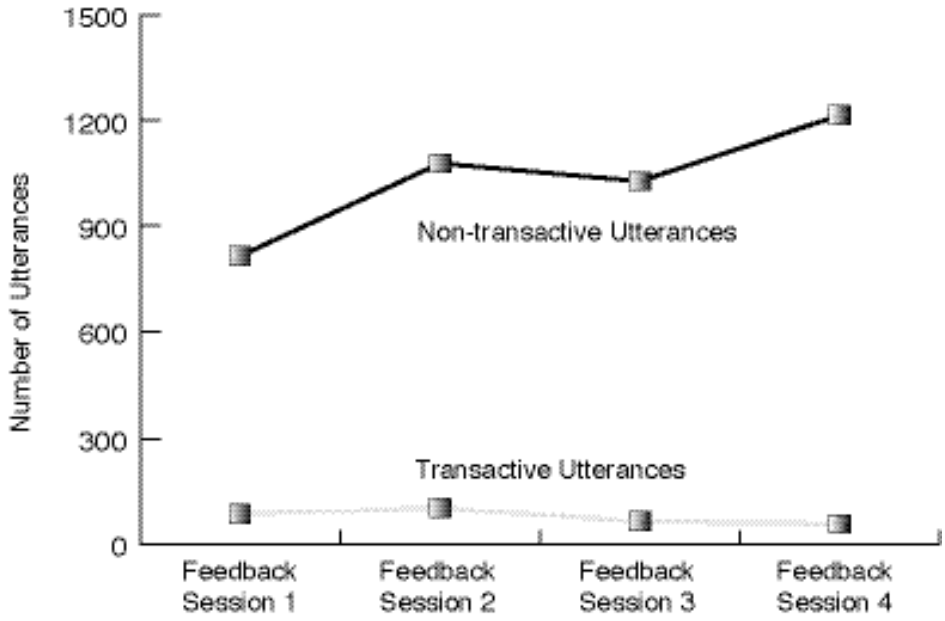
Note. a – Final examination controlled for SAT pretest

Time on Task

A plot of transactive utterances and non-transactive utterances seen across feedback sessions 1 through 4, Table 3, was inspected for the possibility of time-on-task (Cook, 2001b) effects to determine if there was a possible correlation between the number of transacts uttered and the total number of non-transactive utterances made

during the study (see Figure A2). The correlation between the total number of transactive utterances and the total number of non-transactive utterances made across feedback sessions 1 through 4 was not found to be significant ($r = .12$, $p = .29$, *NS*).

Figure A2. Plot of Transactive and Non-transactive Utterances



However, in considering only feedback sessions 1 through 3, a significant correlation was found ($r = .52$, $p < .01$) between the number of transactive utterances and the number of non-transactive utterances. Additionally, in considering feedback session 2, a significant correlation was found ($r = .50$, $p < .01$) between the number of transactive utterances and the number of non-transactive utterances.

Table A9 recapitulates the results of the ancillary analyses. The results are presented by the type of analysis and the outcome of the analysis.

Table A9

Summary of Ancillary Analyses

Type of Analysis	Results
Repeated Measures	Post hoc comparisons using Tukey's procedure found a significant difference in the number of statement transacts between feedback sessions 1 and 2, feedback sessions 1 and 4, feedback sessions 2 and 3, feedback sessions 2 and 4, and feedback sessions 3 and 4.
Partial Correlations	The difference in the number of statement transacts and performance on the NewWorm posttest was not significant. In addition, there does not appear to be a relationship between the number of statement transacts and performance on the SAT posttest. There appears to be a significant relationship between statement transacts and performance on the final examination when controlled for either the NewWorm pretest or the SAT pretest.
Hierarchical multiple regression equations describing feedback session 2	Two of eight equations were significant and described the effect of transactive utterances on the NewWorm posttest when controlled for the NewWorm pretest. In the first equation total transacts was not significant as a predictor variable. In the second equation the NewWorm pretest as well as statement transacts were not significant as predictor variables. Three of eight equations were significant and described the effect of transactive utterances on the final examination when controlled for the NewWorm pretest. In the first equation total transacts was a significant predictor variable. In the second equation statement transacts was a significant predictor variable. In the third equation self-oriented statement transacts was a significant predictor variable. Three of eight equations were significant and described the effect of transactive utterances on the final examination when controlled for the SAT pretest. In the first equation total transacts was a significant predictor

	variable. In the second equation statement transacts was a significant predictor variable. In the third equation self-oriented statement transacts was a significant predictor variable.
Hierarchical multiple regression equations across all feedback sessions	Two of ten equations were significant and described the effect of transactive utterances on the NewWorm posttest when controlled for the NewWorm pretest. In the first equation self-oriented transacts was a significant predictor variable. In the second equation self-oriented statement transacts was a significant predictor variable. Four of ten equations were significant and described the effect of transactive utterances on the final examination when controlled for the NewWorm pretest. In the first equation total transacts was a significant predictor variable. In the second equation statement transacts was a significant predictor variable. In the third equation self-oriented statement transacts was a significant predictor variable. In the fourth equation other-oriented statement transacts was a significant predictor variable. Four of ten equations were significant and described the effect of transactive utterances on the final examination when controlled for the SAT pretest. In the first equation total transacts was a significant predictor variable. In the second equation statement transacts was a significant predictor variable. In the third equation self-oriented statement transacts was a significant predictor variable. In the fourth equation other-oriented statement transacts was a significant predictor variable.
Time on Task	Correlation between the total number of transacts and the total number of non-transactive utterances made across feedback sessions 1 through 4 was not found to be significant. Correlation between the total number of transacts and the total number of non-transactive utterances made across feedback sessions 1 through 3 was found to be significant. Correlation between the total number of transacts and the total number of non-transactive utterances made across feedback session 2 was found to be significant.

Discussion

A repeated measures analysis of variance determined that the statistical difference in the number of statement transacts across all feedback sessions was not significant. The analysis was undertaken outside the scope of the original research questions when it was noted that there appeared to be more self-oriented statement transacts and other-oriented statement transacts than other types of transactive statements. These measures were collapsed together as *statement transacts* in order to further investigate their collective effect on outcome measures.

Noting again that statement transacts are those which operate on a partner's reasoning or are made by the participant in an effort to better understand their own reasoning, it was expected that there would be many statement transacts and that they would persist across the course of the study. However, although more statement transacts were found than any other type reported in this study, as with the other forms of transactive utterances, after an increase in the number of transacts in the second feedback session, there was a general decline in the number of transacts.

Hierarchical Multiple Regressions

The number of statement transacts uttered by participants served as a strong predictor of performance on the final examination whether controlled for the NewWorm or SAT pretests, accounting for between 13% to 41% of the variance during either the second feedback session or across all sessions. It is notable that during the second feedback session the number of self-oriented statement transacts served as a strong predictor of performance on the final examination when the effect of performance on the SAT was considered, accounting for 42% of the variance. However, when considered across all feedback sessions, the number of statement transacts (consisting of both self-oriented

statement transacts and other-oriented statement transacts) uttered during the study was more predictive of performance. These findings suggest the importance of not only the knowledge students bring with them to the classroom as reflected in pretest assessment performance, but also the process of discourse through which participants interact.

That other-oriented transactive discussion of some nature did not perform a larger predictive role was unexpected. One reason other-oriented transactive discussion did not perform a greater predictive role might be the composition of the participant groups. This seems supported when noting that previous researchers (Azmitia & Montgomery, 1993; Kruger, 1992; Kruger & Tomasello, 1986) have noted a connection between group composition and the type of discussion it produces. Kruger & Tomasello (1986) observed the importance of group composition and member relationship in noting that peers are more likely to use transactional discussion than when interacting with someone who is not a peer. Additionally, Kruger (1992) noted that groups with dissimilar peer pairing between members created fewer other-oriented statements than similar peer parings, suggesting that such groups would produce more self-oriented statements. This is further supported by Azmitia and Montgomery's (1993) observation of higher levels of other-oriented transacts produced by friends which could be related to higher levels of within-group transactional conflict and improved problem solving accuracy (Doise et al., 1976). As a result, if the members of the groups had been close friends, we might expect higher levels of other-oriented statements and improved performance on the outcome measures.

The higher prevalence of statement transacts may also be seen as a direct result of the GenScope Assessment Project, which was designed to foster assessment conversations (Hickey et al, 2002). In that the participants were engaged in presentation, improvement, evaluation, and the elaboration of ideas, their use of statement transacts

would be expected (Azmitia & Montgomery, 1993). Moreover, noting that statement transacts represented only 4.65% of observed utterances, their apparent predominance in expressed transacts seems to agree with Kruger's (1992) and Teasley's (1995) observation that the type of discourse is more important than the quantity of discourse.