

# What Is Not Institutionally Visible Does Not Count: The Problem of Making Activity Assessable, Accountable, and Plannable

**Charles Bazerman**

Department of Education  
University of California, Santa Barbara  
Santa Barbara, California 93106  
[bazerman@education.ucsb.edu](mailto:bazerman@education.ucsb.edu)  
<http://www.education.ucsb.edu/~bazerman/>

## Abstract

This hypertext examines from an activity theory perspective the vexed problem of assessment and its relation to planning, accountability, curriculum, and learning. Assessment although only part of the educational process has implications for almost all of education. Local, state, and federal policies that have put great weight and high stakes on a battery of assessment tools that stand outside the daily life of the classroom but are intended to hold classrooms, teachers, and schools accountable for results.

While situated evaluation is an aspect of most human practices, institution-wide testing creates substantial difficulties for the local practices of each class, and particularly creates tensions between student-centered classroom practice and subject-centered expectations. Such tensions have been a continuing puzzle for progressive education. Dewey and his followers regularly preferred to keep evaluation and decision-making local, but for various institutional reasons had to seek larger ways of assessing student achievement without ever being able to develop fully appropriate assessment tools. The teaching of writing has faced a similar dilemma, with standardized forms of writing assessment setting reductionist definitions and expectations of writing, and not directing students towards the highest levels of accomplishment. This study seeks considers genre and activity analysis as the basis for defining and assessing writing tasks through analysis of materials collected from a complex sequence of social studies writing assignments on the Maya from a sixth grade class.

## Contents

[1. Overview](#)

[2. A Personal Preface](#)

[3. The Tension Between Student-Centered and Subject-Centered Curricula and Assessment](#)

[4. Interactions, People, and Large Institutions: The Paradoxes of Assessment](#)

[5. Dewey and His Colleagues Confront Assessment](#)

[6. The Dilemma of Assessment as Seen in the Teaching of Writing](#)

[7. Alternative Assessment](#)

[8. The Need for a New Approach to Assessment](#)

[9. Genre and Activity Theory](#)[10. An Activity Analysis of Planning, Accountability, and Assessment of a Multi-disciplinary Project Based Learning Experience](#)[11. Discussion of the Case Study](#)[12. Importance of Issue and Implications of This Line of Solution](#)[References](#)

## **1. Overview**

This hypertext is an attempt to understand from an activity theory perspective the vexed problem of assessment and its relation to planning, accountability, curriculum, and learning. Assessment has been continually problematic because, although it is only one part of the educational process, it has implications for almost all of education in the way it appears to pass judgment on students, teachers, classrooms, and institutions. These implications have been heightened by local, state, and federal policies that have put great weight and high stakes on a battery of assessment tools that stand outside the daily life of the classroom but are intended to hold classrooms, teachers, and schools accountable for results. Current policy initiatives elevate large scale testing into the dominant, if not sole tool, for gathering information and understanding schools.

This hypertext is also an experiment to see if activity theory may itself provide a tool for developing more challenging, situationally relevant, but generalizable and comparable broad-scale assessment. That is, I wish to explore whether the idea that all local activities are constructed out of typified elements, particularly genres, allows us to compare meaningful situated performances across many different settings. Genres and activity systems, I believe, can help guide us through the dilemmas of local learning and large-scale assessment, just as they guide us through the conundrum of living our improvised local lives with some sense of order, expectation, and relevant skill.

I am particularly interested here in writing as a site of assessment for a variety of knowledges and skills. Almost all assessment, particularly within schooling as in other bureaucratized institutions, requires some form of inscription—even if only a check mark on a form by an assessor watching a student accomplish a gymnastic task. But academic assessment typically involves the person being assessed to inscribe some choice—true-false, a filled-in bubble on a score sheet for multiple choice, or a more extensive and open statement. Of course the more one gives students space to express their thought and knowledge the more control students have control of the discursive space—that is they need to write more in a more open-ended format. This puts higher and more complex demands on their abilities to formulate a response within the situation and allows them to display more talents including those around complex decision-making about relevance of response, developing reasoned relationships among statements, and

selecting ideas and resources to represent in the text. However, as skills, decisions, and products displayed in open-ended tasks become more complex, so do the tasks of evaluating them.

I am also interested in how assessment enters into writing and the impact of assessment regimes on writing and writing education and the use of writing in all forms of education. How people are assessed, how writing instruction is assessed reflects what is valued and evaluated in the competence-- even when “the writing” itself is not putatively the thing expressed, but only “the knowledge” or “the thought” expressed in the writing. Indeed, in such cases where writing is removed from focal attention we have identified the hidden curriculum imperative of writing throughout the academy, across the curriculum.

If writing is indeed, as the articles in this collection all argue, a crucial site at which we construct ourselves and societies, the assessment of writing is the site at which we examine, reflect upon and evaluate who we are becoming as individuals and societies. As writing then serves as an account of who we are, it becomes the place at which individuals, groups and institutions become accountable to evaluation. This evaluation in turn constrains, directs, rewards and punishes particular lines of development, affecting our planning for who we become in our writing and who our students become within our curriculum.

Because issues of assessment are so complex and unsettled--with so much history, so many interested parties, and so many dimensions—I do not find myself able at this juncture to offer a coherent, linear analysis of the issues, problems, and solutions. I have tried to give some shape and substance into my inquiry by two sorts of research. First I follow the history of how Dewey and his followers (variously characterized as progressive or child-centered) dealt with assessment issues. Secondly, I present ethnographic data on a complex of activities in a fifth-grade classroom that reveal something about how the organization of tasks and assigned genres are related to the knowledge and skills displayed in various documents. These sites of inquiry have sharpened the issues, but neither provides the magic solution.

Yet I still feel that an activity approach offers significant help in understanding the issues and pointing to potential lines of solution. Consequently I am putting my thoughts together in a series of small essays or sorties into the issue in ways that are more or less loosely connected. They may be read sequentially, individually, or by following internal links.

Because this issue is clearly not at an end, I will also be including a forum for people to comment on aspects of my presentation and then to continue the discussion.

## **2. A Personal Preface**

As a teacher and researcher of writing, I have long resisted devoting much time and attention to assessment. Assessment, as a separable concern apart from accomplishing an immediate task, has seemed to me to be peripheral to the main business of writing. The entire leverage for improvement, instruction, and purposeful, growth-inducing practice, seemed to me to lie in addressing a communicative task at hand. A person works hard at writing, developing practical

skills, in order to do those communicative things that writing does. Within motivating situations, accomplishing a locally relevant purpose becomes an immediate intrinsic reward. Ultimately the most powerful assessment is how well the writing has worked in the situation and the most important assessment is how one can make writing work better in the future.

Over the years I have held that in order to improve writing we need to understand in fact what writing does within the situations and interactions it functions within, and the means by which it accomplishes or fails to accomplish those tasks. This functional, pragmatic approach has guided my research, theory, and pedagogic practice. In providing guidance to students and feedback on their drafts and final copies, I focused on what students wanted to accomplish in situations and how texts could accomplish it rather than formal performance judged against a general, formal standard, or even a genre-based standard. Genre could provide parameters for defining the rhetorical problem and possible communicative solutions, but could not speak specifically to the effectiveness of any particular text.

Assessment, on the other hand, I understood as something apart from writing. In seeing them as distinct, I was following the pattern set in the literature, which defined assessment largely as an institutional issue (see, for example, Cooper & Odell, 1977; White, Lutz & Kamusikiri, 1997; Wolcott, 1998). Of course institutions had good institutional reasons for assessment, but I was happy to let others more bureaucratically minded worry about that--and I felt we should not let the field be driven by bureaucratic concerns. Given the important institutional roles of composition in higher education in the United States, I appreciated why so many in composition studies seemed concerned by assessment but was troubled that the amount of disciplinary attention it drew and certainly had no desire to contribute to it. I also understood how the testing that was in place had enormous consequences for students and curriculum, focusing the attention of both on testing demands. During my two decades at City University of New York I witnessed and viscerally felt the great strains the Writing Assessment Test (which served as both a placement exam and a requirement for upper division status) placed upon the basic and ESL writers in my classes. I also struggled with developing a curriculum for my own students and for the department that would meet the demands of this exam.

Another layer of my resistance to devoting much attention to assessment came from my own history as a student. Throughout my school career, from primary grade IQ tests through graduate school entry exams, I was very successful at standardized tests, especially the multiple-choice sort. I consistently got top scores, and on the bases of these scores won awards and entry into special programs, was admitted to elite educational institutions. My ability to do well at these tests and my consequent participation in elite programs became part of my identity and personal sense of self-worth, which included distinguishing myself from my classmates.

Gradually, however, I became aware that these tests were not measuring my real learning or accomplishment, and may have even been distracting my attention and commitment from more important forms of learning and accomplishment. As I matured, I began to realize that being better at tests was a pretty slim reed on which to build an identity and a life. While I was grateful for the opportunities success at tests brought me, I started to see that good test results hardly

made a difference in the meaning or value of my life. Further, it eventually penetrated my self-justifying haze that while my test results granted me privileges and experiences, they denied privileges and experiences to others who would have also benefited from the experiences. Then as a teacher I started to understand the self-fulfilling nature of expectations, tests, and experiences for my students. I found that students could perform far beyond previous expectations if you engaged them at a different level and provided pathways for growth. Most assessment measures, however, simply reinforced the student's current position. Particularly for those testing poorly, assessment measures serve as blunt instruments of exclusion, keeping students from just those engaging experiences that might make learning meaningful.

I recognize, however, that we can never escape assessment, which is a central part of any reflective practice. Every time I write, revise, or just choose words, I assess the effectiveness of my emerging text. And every time I put comments on a student's paper, or give writing advice, or even discuss a student's possible choices in writing, I engage in assessment. Assessment is inherent in any performance, teaching, or learning task, because writer, teacher, and student are focused on improving knowledge, understanding and performance. Such situated assessment of performance in process is particularly strong in craft-like tasks like writing or music performance that allow for improvement and choice-making based on self-monitoring and reflection. Self-monitoring, reflection, and consequent modification of activity has been from the time of Vygotsky (1986) a key construct of many learning and performance theories (see, for example, Bruner, 1990; Schön, 1983; Giddens, 1984; Hillocks, 1995). As a result I have seen these literatures of assessment in performance as being of much greater interest and importance than the literature on institutional assessment. I have also found useful and interesting the literature developing tools for responding individually to student writing (see, for example, Anson, 1989; Lawson, Sterr & Winterowd, 1990; Straub, 2000) (*see [section 7](#)*).

From the point of view of writing and the teaching of writing, I have seen for the most part little necessity for assessment beyond the actual performance of the learning tasks and production of the text objects. Writing leaves its own record of success and failure, competence demonstrated and things to be worked on. Further the teaching of writing through conferences, group discussion, teacher feedback, peer editing, revisions and other processes of current pedagogy provided so much situated and task-relevant assessment that there seemed little purpose for further formal assessment except for institutional accountability and placement. Even the standard university course assessment tools of mid-term and final exams seemed redundant or inappropriate when we had so many pieces of commented-on work throughout the term to track in great detail what the students have learned and accomplished. The practices now labeled performance assessment and portfolio assessment (*see [section 6](#)*) in educational textbooks on assessment have a long history in the teaching of writing. One can in fact plausibly argue that the current wide popularity of alternative assessment in education grew out of teaching of writing practices in the United States, particularly as disseminated by the regional and national writing projects over the last quarter-century.

In my teaching and writing for myself, accordingly, I have developed an increasingly strong sense of what I want my writing and my students' writing to accomplish. These expectations

provide the basis for practical, task-oriented assessment activated at every stage of the writing process--from early discussion of students' ideas about how to approach the assignment, through responses to drafts, to final comments on what has been achieved. These same expectations shape other directive comments supporting the assignment that are not in direct evaluation of or response to a student production. Together, comments and production-directed assessments in the midst of the writing process anticipate the emergence of writing that speaks forcefully to the task at hand. The aim of these proleptical comments is not to produce formal assessments of student work, but to help students develop their own internal standards and modes of self-assessment as part of skilled writing competence. I want them to develop their own critical "writer's eye" for their own work to make them stronger, more effective writers. As they develop a sharper eye for their own work, I can then increase the demands for even higher levels of performance, through my evaluating their emerging work against more subtle task-based criteria.

This kind of *in situ* evaluation, however, is labor-intensive individualized work, sensitive to the student's own level of understanding, the particulars of the task, and the approach the student has taken toward this particular writing problem. I am constantly trying to determine what the student is trying to do with their piece of writing, directing that approach to make it useful and appropriate for the task at hand, and then demanding they do it even more skillfully. And I must do this all within terms that speak to the student's current level of understanding and perception of writing.

This individualized task-focused increasing level of demand, supplemented by tools to meet the demand, provides meaningful ways of increasing the level of challenge. Mass education and the institutional need to provide comparable evaluations of student work, however, are not easily reconciled with this kind of individualized, task oriented *in situ* development. Consequently, within institutional assessment it is often difficult to identify higher levels of accomplishment and challenge that are meaningful and motivating to students. The alternative is to fall back on atomized, alienating, disengaged formulations of the competence to be structured as curriculum and then to be assessed as atomized skills (*see section 3*).

Because accountability to higher standards has traditionally meant an atomization of abstracted skills and formal appearances that are not sensitive to the expressive, creative, rhetorical, and/or intellectual workings of texts and writers, many experienced teachers of writing have similar qualms about grades and other institutionalized assessment (Allison, Bryant, & Hourigan, 1997; Zak & Weaver, 1998). Even when criteria are given some disciplinary or other situational specificity, such assessment criteria are often viewed as inimical to the student-centered classroom. By extension, assessment standards drawn from traditional disciplines are viewed as distancing and alienating to students trying to discover their own meanings in their writing that will have force and motivated purpose. Thus writing pedagogy since the 60's (see, for example, Rohman & Wliecke, 1964; Macrorie, 1968, 1970; Graves, 1978; Murray, 1968; Fulwiler, 1987) with some justification has reached back into what students have most at hand and about which they have the most developed stances—that is, their own histories, experiences, commitments, feelings, and self-defined issues. Such pedagogies teach students many things--but not

necessarily what students need at the university. By having students look inward to make the most of what they bring with them, these pedagogies provide less encouragement to look outward to what the university has to offer in the form of resources and scenes of new interaction. Even when such pedagogies bring the students into a vibrant life of writing, they can distance the students from the remarkable and highly motivated epistemic activity of the modern university.

The disciplinary and epistemic culture of modern universities does limit strongly the role of standardized assessment, as universities prize individual contribution and thought. Ideological commitment to individual student growth and the responsiveness of instructors to students whom they find have individual accomplishment or promise often softens the massification and standardization of even the most factory-like campuses and truly defines the culture of the more individualized campuses. Further, the American higher education system of majors and electives keeps most assessment local to grading for individual courses. Institution-wide assessments tend to be used only at the earlier curricular steps (for example placement exams or exit from required first-level courses) and are perceived as guaranteeing only minimal performance, well below the expectations and hopes for most students.

Since I have moved from an English department, where my attention was directed mostly at higher education and beyond, to an Education department, I have had to take more seriously the ways in which testing regimes impact elementary and secondary education. I can no longer remain obtuse about issues that the freedom, individuality, and faculty prerogative of university life have protected university teachers of writing from. Teachers and students in K-12 schooling are not nearly so well buffered against pressing demands for large-scale assessments. State and local governments that pay for education, the national government that has made educational competitiveness a priority, policy-makers at all levels, administrators who want to maximize teacher and student performance, and parents who want assurances about the accomplishment of their local schools and their own children—all are constantly at the classroom door asking for some accounting of achievement. Further, educational policy-makers and curricular designers need to assess the skills and knowledge of students in order to provide appropriate instruction, support, and challenges so that students may reach higher levels of more ambitious disciplined tasks. Policy-makers also need to be able to assess the effectiveness of forms of instruction and instructional tools to help make institutional choices that will affect the lives and educational experiences of many children (*see [section 4](#)*).

So, I remain ambivalent about assessment, recognizing the imperatives and the dangers. I am wary about what they measure and what they encourage; at the same time I believe it is better to know something than nothing about what people achieve under what circumstances and with what supports. I remain troubled that the detailed *in situ* assessment that is part of intelligent action gets obscured in weak measures spread over large numbers removed from contexts of meaningful action.

As I have become engaged in this project I have become ever more aware of how much thought and experience of practitioners, policy-makers, and researchers have gone into the realms of

curriculum planning and assessment. I am humbled by the long commitment, energy, work practice that has entered into the accumulated wisdom, as well as the practical working out of the complexities of the world that have been factored into practice. And I hesitate about having the presumption to add anything useful to this discussion. My fragmentary comments here are only meant to explore whether my ambivalence, which perhaps reflects both sides of widely held divisions on the subject, might provide some different ways of looking at the issues and paths for solution. I also hope that the particular theoretical, methodological, and descriptive tools of genre and activity theory (see [section 9](#)) can provide some help in creating institutionally visible and credible ways of organizing and monitoring the inevitably individual process by which learning happens and people grow (see [section 12](#)). In a democracy there is no satisfactory alternative than to make each person an intelligent articulate assessor of his or her life situation, an eloquent spokesperson for his or her interests, and an effective communicative actor to carry out a successful, contributing life. Any testing regime that undermines this goal of education--no matter how convenient in distinguishing the skilled from the unskilled and in providing accountable goals for achievement for students, teachers, and schools--undermines the intelligence and vitality of our democracy.

### **3. The Tension Between Student-Centered and Subject-Centered Curricula and Assessment**

Traditional skills and knowledge have consistently posed a dilemma for progressive education (see [section 5](#)). Within student-centered, inquiry-based, activity-oriented curricula, how can organized, disciplinary, and socially expected learning be planned, assessed and ensured? This tension, present from the earliest days of Ella Flagg Young and John Dewey's experiments with activity-centered classrooms (Tanner, 1997) and William Heard Kilpatrick's (Kilpatrick, 1951; Tenenbaum, 1951) proposals for project-based curricula, has reemerged in every cycle of child-centered reform. Dewey frequently insisted on the importance of received knowledge and traditional disciplines being brought in relation to motivated student inquiry and life practice (from his early *The School and Society*, 1899, and *The Child and the Curriculum*, 1902, to *Experience in Education*, 1938). He regularly argued that traditional disciplines needed to be integrated into a progressive curriculum (see Russell, 1993). Despite his vociferous arguments on this score, child-centered educators have not found consistent and robust ways to plan, guarantee and display the achievement of traditional disciplinary competence, such that it has regularly remained a foregrounded goal in all schools that affiliate as progressive. Even more, the proponents of child-centered education have never been able to satisfy critics and policy-makers that disciplinary expectations will be consistently met in the child-centered school (see Tanner & Tanner's discussion of the Lincoln School, 1990). This unresolved tension has left child-centered education with little ammunition in the current climate of imposition of standards and testing to assure achievement in traditional areas. The alternative assessment movement (see [section 6](#)) has attempted to address this absence, but has not yet produced the kinds of data that have satisfied those insisting on widespread accountability for achievement in traditional areas of organized knowledge and skills, as expressed in direct testing of content knowledge.

Few parents, politicians, and educators on any side of any fence would deny the value of student motivation, engagement, and curiosity. Nor would many question the educational and motivational value of drawing on the materials of the daily life of children and adults in contemporary society. Nor would many oppose the desire for children to develop into competent citizens, contributors, and social beings in the world they will live in. On the other side of the ledger, few on either side would deny important value is to be found in the accumulated skill and knowledge developed over the millennia of recorded human experience. Nor would they deny that life in contemporary society requires familiarity with a wide range of knowledges, disciplines, and disciplined skills. Indeed, at the juncture of these two views, most parents, politicians, and educators do understand the function of schools to bring the engaged and motivated child into the world of skills, knowledge, and socially organized practices of our current economic, social, and cultural life.

The problem is how to bring these two sets of valued objectives together. The most direct way to ensure student engagement and inquiry is to follow student interests and projects, keeping the curriculum open to capitalize on impulses toward knowledge that teachers identify in the children and to increasingly make students responsible for their own learning and choice-making (as with Kilpatrick's project method). An interest in the resources and accumulated wisdom of humankind might be assumed to follow, perhaps as made relevant and available by the teacher and others who understand the location, value and importance of such cultural resources. After all, these resources, knowledge, ways of creating, practices, facts, and other components of cultural knowledge were developed in response to the processes and challenges of life. Not only is this cultural knowledge useful but it has created the current organization, conditions, and practices of contemporary life. This cultural knowledge is embodied and assumed in the built physical, social and symbolic environment. So to be caught up into any real project will inevitably engage you with the world and the traditions by which we have come to cope with and shape that world.

This has indeed been the position of many involved in progressive education and project-based curricula over the years (see Katz & Chard, 2000; Collings, 1931; McMurry, 1920; Passe, 1996; Erickson, 1995; Rance, 1968; Webster & Smith 1927). However, this individualized inquiry into and discovery of the world puts large burdens on the teacher. The teacher must be in touch with the dynamics of each student's or group's inquiry, knowledgeable in the many domains students might find useful, and resourceful in finding the supports students might need. The teacher must be able to persuade students that they need to pursue these cultural resources and not simply reinvent wheels. To stretch students, teachers must also be persuasive in raising the bar of achievement beyond the spontaneous criteria of adequacy that children might develop on their own. Even if a teacher knew what was going on in each inquiry, could provide relevant resources, and could push students towards more ambitious and refined inquiries, there is no guarantee that students would gain comprehensive and orderly skills and knowledge in any domain. There is no certainty, therefore, that students will be prepared to pursue more advanced disciplinary studies. Indeed it may seem that such an inquiry-based approach keeps secrets from

students about the resources of specialized knowledge until students stumble into what disciplines have ready at hand.

On the other hand, it would seem that the most direct way to ensure that students learned the skills and knowledge we desire them to have is to instruct them directly and then explicitly examine them on that knowledge. A coherent academic curriculum designed by state or national experts in the desired areas of knowledge and practice would then form an orderly sequence for teaching, learning, and testing. Such curricula could provide guidance for teachers in defining what they need to teach and what resources are needed to support the learning. Relevant materials could then be provided systematically by textbooks, library resources, and institutionally provided resources. An orderly curriculum would then define increasingly higher levels of practice and evaluation as students, on the basis of foundational knowledge, are required to address more advanced and subtle materials and tasks. This reasoning characterizes the standards and testing movement: Only well articulated and disciplinarily rigorous standards and tests to measure achievement of those standards can increase the general level of achievement and can ensure that all schools will teach high level disciplinary knowledge and skills. Further only such clearly articulated and accountable performance will guarantee that all children will have access to desired knowledge and will provide equity of opportunity among schools despite economic and demographic variation. Further, some suggest only explicit standards and direct accountability will provide guidance to teachers who are not uniformly well versed in disciplinary knowledge and therefore cannot provide students the flexible, insightful entry into disciplinary required of successful student-based, project oriented teaching (see, for example, Ravitch, 1995; Stotsky, 2000).

A standards- and testing-based curriculum, however, puts high burdens on students for continuing attention, cooperation, and understanding of assigned material that may be far from their immediate interests or concerns. It is also then left up to individual students to see how this material applies to the varied situations that make the knowledge meaningful and valuable. Critics of standards- and testing-based education regularly point to the many students who do not find their way through the world of externally imposed tasks and tests, whether for reasons of personal difference, cultural and class mismatches, or particular life and family difficulties. Critics point to the difficulty of learning through a day of alienating activities with only shallow extrinsic reward of grades. Many students under such conditions may lose motivation and direction. Even students who do succeed in this world of requirements, may only develop a superficial understanding and rote practice, critics note. Only a lucky few may really find the intrinsic values and conceptual understandings that underlay our cultural legacy that makes it continually applicable to life. Further critics point out that extrinsic rewards distort learning and motivation, which then rapidly vanish once external rewards are removed.

In practice, schools are rarely so dichotomous that teachers and students are forced totally into one or another extreme position. Even in the current political enthusiasm for standards, tests, and regulation, while it puts severe pressure on students and teachers to perform well on externally imposed tasks, cannot stop children from expressing their enthusiasms and interests. Further, alert teachers cannot stop noticing and responding to students' expressions of interest,

bafflement, and lack of interest. Rather the tension continues, with individuals feeling the waves of school reform pushing one way or the other. Teachers, students, and schools seek solutions that accommodate to some degree the needs of teachers and students, the demands of organized knowledge and practice, and the realities of student motivation. Even the most traditional schools celebrate teachers renowned for their ability to reach out and motivate students. Even the most student-centered schools make some assessment of students' progress in traditional areas of knowledge and provide some structured planning for helping students who do not seem to be moving into areas of basic learning spontaneously.

An adequate consideration of assessment requires putting together--but also keeping distinct--planning and assessment. Tests influence the planning of educational activities as well as the orientation of teachers and students when in the psychological penumbra of the test—a phenomenon called assessment washback. This is viewed sometimes with delight and sometimes with horror, depending on one's view of the values and orientations embodied in the test. Assessment when used for formative purposes rather than evaluation and accountability, nonetheless still directs attention and curriculum, for formative assessment identifies specific areas of student accomplishment and skill, or lack thereof, to be worked on. If formative and evaluative assessments are so far from what teachers and other curricular experts think must be taught, then the tests appear irrelevant. The formative does not provide useful information for curricular planning, and the student learning within the planned curriculum may not be evidenced in the evaluative tests. The believers in the items measured in formative and evaluative assessment will view the curriculum as misguided and educationally irrelevant, while the people committed to the values of the curriculum will see the tests as an imposition or worse. This situation seems in many ways to characterize the conflicts currently in literacy education.

As is often noted those items most easily tested are not necessarily the most important or challenging educationally, nor do they provide the most motivating and engaging for learners. The process of testing often isolates and anatomizes skills, to provide comparable results across disparate situations and students. Testing of isolated anatomized skills also facilitates developing individualized profiles for each student and planning individualized support based on what they can and cannot display under testing conditions. But complex, high level accomplishments require both integration of multiple skills and some higher order creative, analytic, and synthetic skills that evaporate if disaggregated into atomized tasks. Motivated engagement and creative learning interaction with coparticipants that lead to complex and more advanced learning also are not supported by testing of isolated anatomized knowledge and skill. Thus regimes of testing tend to have a reductionist and debilitating effect on curriculum

For those who hold a strict behaviorist model of learning focused on specific skills motivated by extrinsic reward, atomized, abstracted assessment is not at all a problem. In fact, it is a significant part of the reward structure. And planning sequential introduction and rehearsal of anatomized parts and their sequential complex combination provides guidelines for planning. At the extreme one could see the benefit of this approach for such skills as typing or other tasks that would benefit from automaticity—spelling, math facts and simple operations, geographical and historical facts. But such specific knowledge elements only become meaningful and motivating

in the long term as part of a larger practice where knowledge and skill are used and valued. Even typing and spelling skills are only maintained at high levels when they are put to motivated use, whether to earn money or facilitate one's own writing. Continued learning and engagement with geographical and historical facts occur only within the process of understanding our world and engaging in compelling issues. These kinds of engagements draw on integration of knowledge with imagination, problem solving and purposeful thought. Ultimately learning even the most anatomizable and abstractable knowledge is facilitated by interesting, engaging activities.

From the point of view of a rich complex curriculum, other sorts of assessment may seem preferable for formative curricular planning purposes. Vygotsky's (1978) concept of the zone of proximal development suggests rather than assessing fully accomplished skills you may be better off assessing kinds of activities the student could participate in and learn from with appropriate support. Such assessment would identify those areas where they would know enough to keep engaged and to comprehend new material, but they would not know so much that all would be easy and familiar. In a well-identified zone of proximal development they would need to pay attention to and learn from teachers and other students who could guide them on what they needed to know. Similarly at the end of the learning episodes one could measure what students could do on their own and what more advanced learning activities they could now engage in with others. But zone of proximal development assessments have typically been individualized and approximate, based on a one-on-one encounter between the student and professional, resulting in a qualitative, descriptive account. Not only is this a cost intensive process, it is ill suited to the constraints of large program assessment and curricular planning.

Even if the problem of assessing what students could participate in were solved, the question of what you want them to participate in to foster advanced learning looms. What activities and environment will foster particular kinds of advanced learning? What kinds of advanced learning are you seeking that go beyond easily isolated, anatomized and quantified items that might appear on a standardized exam? What kinds of activities will direct students' attention and excite learning and practice of more advanced skills?

One apparent way out of this dilemma is to believe that once students get beyond the assessable basics, intrinsic task-based motivation will take over, making the tests a barely noticed, irrelevant institutional necessity. Unfortunately that does not appear to be the general case. Once students are caught up in the extrinsic rewards of testing and scores, it seems to have deep and continuing effects, distracting students from intrinsic rewards (Kohn, 1993), and defining the educational process by the extrinsic evaluative record one establishes. Grades appear to become even more important in higher education as students become institutionally savvy (Becker, 1968), despite the apparently greater opportunities for pursuing individual interests, career motivations, and direct engagement with actual social problems.

What is needed is some way to bring assessment together with situated engaging activities. We need assessment and planning tools that serve the needs of institutions as well as of individual learners, that do not put engaging learning experiences in struggle with low-meaning test situations. We need a strong tool for rich curricular planning that does not thin the educational

experience. Genre and activity theory, I will argue in later sections of this text can provide ways to identify the forms of advanced practice that require and put into practice specific domains of knowledge and skill at the same time as engaging students in challenging and rewarding projects and goals. These forms of practice will in turn provide a matrix upon which to assess learning.

## **4. Interactions, People, and Large Institutions: The Paradoxes of Assessment**

In interaction people assess the skills and knowledge of their coparticipants so as to be able to communicate and carry out joint projects successfully (Clark, 1992). In individualized learning interactions the mentor regularly accommodates to the perceived knowledge of the learner and supports the learner at the same time as assessing reachable challenges. In order to provide the most relevant support for learners, teachers assess student knowledge and participation in a variety of informal and formal ways. On the fly as interactions unfold, teachers note child reactions, confusions, hesitations, and quick responses. Sometimes assessments are more organized, perhaps using standardized tasks, though each degree of standardization necessarily obscures some awareness of the particulars of each child's understanding, structures of perception and thinking, and confusions. Further each degree of standardization, by defining tasks in ways that are uniform across situations, removes knowledge and skills from a local motivating task; consequently the focus, response, and motivation of the learner potentially becomes transformed and possibly vitiated. The response becomes a performance to cooperate in a testing situation, with no intrinsic reward for participating in the task itself.

In interaction there are also what we might call exit assessments—recognition that we have done all we can do and thus have no motive for continuing. This is typically a recognition that we have accomplished what we needed to together, or that a coparticipant understands the communication, or can and will do certain things. On the other hand, the recognition can be that the coparticipant is fixed on a course of action in which we cannot or do not wish to intervene further. In either case we find it no longer useful to worry about communicating a specific point or eliciting cooperation on a particular matter. Nor do we then need consider whether our intentions or thoughts should be modified or abandoned given our coparticipant's alternative knowledge, perspective or understanding. That is, we no longer need to accommodate to find a meeting place for our ongoing interaction. We judge, for example, that our visitor understands the directions well enough to drive back to town. Alternatively we may assess that no matter how much we insist that their planned route is not possible, they will go their own way.

In teaching interactions, we assess when to continue or withdraw support, when to extend or cut off an assignment. We judge when more explanation would be extraneous, when we can leave the student alone to work on the problem, when we can confidently move on to the next topic, or when we should just leave off a subject for the day because we can get no further. In a class with many students, however, we have fewer interactions with each individual by which to measure each individual's understanding and particular trajectory of learning. In that case, generalized tools of assessment can help us to decide when the class as a whole does not need more time on

an activity or skill or topic. Further the generalized assessment might help us identify who might need continued individual support. But of course generalized tools standardize what it is to be learned and, even more significantly, the trajectory of learning--what topics are best completed in what order.

Any educational experience that involves more than a few individuals--no matter what the educational goals--would need to confront group assessment issues to be able to organize the work of many people gathered together for the purposes of learning. Even if a classroom were to be organized around individualized learning, some standardized measure would typically be needed to keep account of individual student accomplishment and to identify appropriate next tasks for each student. In large individualized classes, in fact, the pressure for standardized accounting may be all the greater because the co-orientation effects of all interacting in a common group under the leadership of the teacher would be weaker. That is, each student would be going down his or her own path with less time orienting to the group and teacher. While the teacher might have more individualized time with each student than in a common lesson, each student would have less total time together with the teacher and classmates, engaged in common activities. Further, at each intermittent individualized meeting with each student the teacher must quickly orient to the student's current state of work and learning.

Thus assessment is closely tied to the organization, planning, and constant unfolding of activities and is not just an after-the-fact evaluation of student accomplishment which may have consequences for rewards and punishments. As long as assessment is carried out in a spontaneous, interactional mode--while having all the naturalness, motivation, and creativity of the interaction in process, it is organically integrated with the immediate motivated goals and activities at hand. Insofar as assessment becomes an activity in itself, becomes standardized to allow aggregation or comparison, or otherwise becomes perceivably separated from locally motivated activities, it loses the spontaneous and motivated engagement as well as local relevance. However, we do have good reasons to want to abstract and standardize assessment tasks and measures.

Among those good reasons are the very reasons schools were established and maintained. Spontaneous learning interaction inevitably happens every day as we meet the challenges of living in the world. All children, whether in or out of the school learn rapidly the practices of the communities they spend their time in. In some societies these spontaneous interactions then become more organized and focused in an individualized apprenticeship system. Some societies organize children's learning more formally, and more removed from daily life, in schools. Societies that organize schools and particularly require universal schooling, recognize that the community has certain historically emerged bodies of knowledge and practice that all participants in the society need and should have access to. The members of the society perceive these skills and knowledges are so important that they cannot be left to random interaction. The knowledges and practices for which schools are given the responsibility are typically perceived as complex and organized, such that random, spontaneous encounters will not provide one with a full, extensive and coherent enough familiarity with them to meet the needs of living in that society. Indeed, in modern globalized society, one must spend a large fraction of one's life in

orderly, somewhat sequential, institutionalized learning to be a competent practitioner in most of the crafts of life--especially the crafts which provide prestige, power, wealth, and socially valuable work. Schools, academies, universities, research institutions, textbooks, encyclopedias, handbooks, professional societies, scholarly journals, professions and professional accreditation are all parts of overlapping institutions, artifacts, and practices. In the middle of the 20th century in the United States, graduation from high school became a near universal necessity, and now substantial higher education appears to be prerequisite to an economically viable life. Graduate training is becoming an expected part of most professional or leadership careers. Careers built on extra-curricular crafts--such as sports, entertainment, and traffic in illicit goods--do provide attractive but risky alternatives. These alternatives, however, carry enormous weights of constant intrinsic assessment. The high risks are precisely because winners and losers, those judged to have talent, skill, or luck and those without, are so effectively and brutally distinguished.

Because today few reliable alternatives to schooling provide access to successful lives, we need to be able to make success in schooling accessible to everyone. It is imperative students encounter areas of complex knowledge and endeavor without getting in over their heads, going down a faulty path, having major gaps in knowledge, developing bad practices, winding up getting lost along the way, or losing motivation. We need to be able to understand what students can and can't do in those domains we deem necessary so as to facilitate continuing progress. We need to know where along these paths students are so that we can help them to comprehensive and skilled competence in many areas, many of which build on prior competences and knowledge. Then ultimately we need to know whether students have gained adequate competence in relevant areas for them to enter into careers and areas of practice that require specific high-level knowledges and skills. We need to be able to assess whether individuals are knowledgeable enough to be credentialed as teachers or doctors or accountants.

In addition to the assessments immediately applicable to individual and group learning, assessment of learning, teaching and other school activities are needed to organize large educational institutions. Systemic planning, allocation of resources, and adjustment of the system require means of monitoring what is going on in many sites. This need for aggregating information occurs at the level of school, district, state, and nation. Some policy planning even is aided by international information. The movement towards standardization, quantification, and testing for individual and institutional accountability has become ever more pressing as institutions and areas of planning have grown ever wider. As is well known, the needs of the military for both assignment of individuals and training provided the starting place for mass testing. Military training standards have often been the leader in curriculum standardization, because of the need for absolute reliability in specific forms of knowledge and practice, and the frequent reassignment of individuals that creates the need for replacements with appropriate competence to step immediately into work groups. Brecht's drama *Mann ist Mann* sharply and critically exposes the loss of identity and personal affiliation that is the ultimate consequence of full standardization and replaceability. Nonetheless, educational institutions have some need to assure the levels and particulars of learning as well as to provide appropriate learning situations for various students. To assure the level of education within states, states have moved toward

centralized curriculum planning and/ or standards, tied to specific subject performance and testing. In the post World War II era, the rapid increase of numbers of students in higher education and the increasing size of institutions has led to great reliance on standardized assessment measures for admissions and placement, to sort through students of varied educational experience within many different school systems. And in the last half-century, United Nations agencies and other international organizations concerned with development throughout the world, have led to an entirely new level of monitoring and assessment in order to help each nation improve its own educational system.

Those who provide the resources to maintain these systems do have a justifiable claim on knowing how well each part seems to be accomplishing its task, in order to monitor the flow of resources. Thus, although politics is often rightfully viewed as an imposition on educational practices, nonetheless assessment of educational practices seems to have an inevitable and warranted place within government and politics at all levels of jurisdiction. Even courts must have means to ensure the accomplishment of mandated governmental responsibilities toward education, equality of opportunity, and rectification of inequities.

Finally assessment has some justifiable role in assuring that people are working hard at the tasks the schools are dedicated to. High standards and expectations are important in setting and reaching institutional aims as well as individual goals. This desire to raise meaningful expectations and make visible how well teachers, principals as well as students are responding to them can justify substantial monitoring on the state and federal level. It is also a key rationale of the standards movement.

Unfortunately, these systems of monitoring seem regularly associated with competition and mistrust. Monitoring means someone is looking over your shoulder and that implies they don't trust you to do the right thing on your own. Assessment makes one person's or group's accomplishment, skill, or other characteristic visible to someone else, and that interest in viewing others work and characteristics almost always has a scent of mistrust. That scent of mistrust can become a very strong odor if there is any indication that people are being monitored because they are not toeing the line, achieving enough, working hard enough. Sometimes that mistrust is truly warranted as malfeasance, corruption, incompetence, discrimination, lack of ambition, or other dysfunctions have at times overtaken schools. Systems of assessment serve to expose deficit and assign blame for those deficits. If results do not meet expectation, it is easy to think there is someone or something at fault--someone is not doing their job, whether students, families, teachers, administrators, or the system. If there is major dysfunction such deficit needs to be exposed, but it is in the logic of assessment systems to generate the appearance of deficit even if there is no clear dysfunction, for someone or some school has to turn up at the bottom. Standardized assessment creates opportunities for comparison and unequal reward. Doing well on an assessment is a way to earn respect, dispel distrust, and direct attention to those who do not do so well. The issues of distrust and competition can become all the greater when assessments are abstracted and quantized, and when the results are viewed from a great distance institutional and or geographic distance, as in a state legislature far from the individual students, teachers, and classrooms.

Because assessments often grow in climates of distrust and competition and then support increasingly distrustful, competitive atmospheres, they are likely to tear at the fabric of the local motivated activity, spontaneous inquiry, and personal knowledge building. Standardization and abstraction of assessment seem incongruent with the personally engaged interaction within which people learn at a high level, the processes by which individuals become competent, respected, knowledgeable, participating members of communities. The abstractions of school, especially the larger institutional abstractions, though paradoxically part of the underlying need for schooling and part of the history of growth and massification of schooling, remove us from the processes by which individuals are engaged and learn.

Dewey among others counseled sensitivity towards the way individuals act and develop within social interactions and settings. He also, however, strongly believed in the value of the accumulated wisdom and experience of the community in providing tools for addressing the practical problems of living and in creating a rich environment and resource for each individual's development. He also understood that students were not individuals but members of immediate and larger communities in which they developed identities and toward which they learned responsibilities and made contributions. Since he and his followers have been central to the development of American progressive, child-centered education, they sit right in the middle of the paradoxes of assessment. Their struggles with assessment, which I will explore in the next section, can help us frame the problem of what we need to demand from assessment so it will truly serve education and not undermine it.

## 5. Dewey and His Colleagues Confront Assessment

John Dewey, from his earliest publications on education, understood that disciplinary knowledge was not an imposition on the child, but the very means of participating in society, and even more the vehicle of personal development. In the 1899 lectures on the University Elementary School that were to be published as *School and Society*, he notes “under present conditions, all activity, to be successful, has to be directed somewhere and somehow by the scientific expert—it is a case of applied science” (1900, p. 23). This disciplined application within the occupations of life

supplies the child with a genuine motive; it gives him an experience at first hand; it brings him into contact with realities. It does all this, but in addition it is liberalized throughout by translation into its historic and social values and scientific equivalencies. With the growth of the child's mind in power and knowledge it ceases to be a pleasant occupation merely and becomes more and more a medium, an instrument, an organ of understanding—and is thereby transformed. (p. 22)

By providing the means of expanded understanding, disciplinary and disciplined knowledge becomes incorporated into who the child becomes, what the child can think about, what problems the child can work on, what roles and opportunities in life the child can take on.

Within the University School the curriculum was built around this intersection of the child's interests and the way organized knowledge can serve those issues—brought together around

specific real or realistic problems. Thus child development was structured within the motivational force of projects and problems and within the intellectual resources of organized knowledge brought to bear on each of the activities. Given the democratic and participatory principles of its founding, the organization and practices of the school depended very much on both the staff and children who came together during the school's brief life from 1896 until 1904. Ella Flagg Young, an experienced teacher who was Supervisor of Instruction, has been particularly recognized by Dewey and others as a leading light in the practical development of the school curriculum. Laurel Tanner's history of the school, based on the archives of the school, reveals for example that lesson plans and reports were written from the double perspectives of "the Child's Standpoint" and "the Teacher's Standpoint." (1997, p. 47). This encouraged all lessons to be thought of as meaningful from the student's point of view as well as to engage substantive disciplinary learning from the teacher's point of view. Actual historical and current problems, such as water management for Chicago or agricultural production or the manufacturing of textiles, gave shape and substance to motivating challenges.

The school favored an integrated approach to knowledge brought to bear on problems. French, for example, was taught in the context of cooking and mathematics in relation to textile manufactures (pp. 76-77). Nonetheless, disciplinary knowledges were such a significant part of the problem solutions and development of students and faculty that Dewey soon insisted the school adopt a departmental form of organization (p. 98). Yet it was in practice, and not the abstract knowledge of disciplines, that assessment of learning and ideas would come, as would the pressure for disciplinary learning and disciplining of development. Dewey praised the concreteness ideas gained from embodiment and testing in practice. Students working on Chicago water management would not be assessed by a test of skills displayed, but by a comparison of the class's solutions to those historically adopted by the city. The results of their mathematics would be in the success of their calculations to carry forward their experiments in manufacturing.

Within the University Elementary School, accountability, planning, and assessment of success remained local, within each classroom. Planning and evaluation of results ultimately rested on the individual teacher, although there was constant group consultation and support. The greatest force for assessment was the internal dynamics of the projects students engaged in. Did they know, learn enough to make the projects succeed? Were the practical problems solved? And did the teacher see any immediate opportunities to expand the learning appropriately? The learned knowledge was aggregated and accounted in the lesson plans, but there was no external measure placed on the learning.

This local system was based on trust, intensive consultation, and collaboration within the school, flexibility in the curriculum, and resourcefulness of the teachers. It also depended on the protection of the university, the status of the school's experimental nature under university supervision, and its voluntary student population drawn largely from university-affiliated families. It was not held accountable to public financing and policy, responsibilities to a larger system, or inspection by various concerned public groups. Further it thrived over a century ago, before the weight of specific college preparation and career credentialing became so well-

defined, compulsory, and discipline-based. While the curricular planning and assessment procedures at the University School seem to present an ideal for one side of the equation, they did not deal with the whole picture which any school must now necessarily address.

The idea of discipline being natural to activities and projects was elaborated in George Coe's small volume on *Law and Freedom in the School*. To counter the standard view that law stands in opposition to will and freely chosen purposes, Coe argues that projects are in fact the law of nature, the way we organize our ambitions and activities, and that further they are naturally disciplining. Projects force us to confront the realities of nature and thereby discipline our understanding and ambitions. Additionally, teachers gain real authority by submitting to the demands of the project alongside the students. Finally projects provide the motive and means for self-government and the development of a sense of moral law.

While Coe again expresses the faith that engagement with the real challenges of projects will provide criteria of success, naturally disciplining processes, and motive to engage seriously with all forms of useful human knowledge, he provides no concrete mechanisms. Further, discipline is treated so intrinsically that any external inspection or reporting invite the undesired dynamics of compulsion. Externally established demand runs counter to the organic discipline at the heart of his account.

In placing all discipline and drive for knowledge within the organic dynamics of confronting reality, Coe is consistent with the great champion of the project method, William Heard Kilpatrick. In the writings of Kilpatrick and his followers there is little concrete way to confirm or monitor student learning beyond description of the specifics of each idiosyncratically chosen and uniquely pursued project. Nor is there much in the way of general guidance in planning project learning activities beyond description of exemplar projects.

One striking exception is Ellsworth Collings's experiment aimed at proving the value of the project method. In 1924 Collings published a four-year study of a one-teacher rural elementary school in McDonald County Missouri, organized on the project method. Two other neighboring schools with similar characteristics in the same county, but retaining a traditional curriculum, were used as controls for the assessment measures. The three schools had between 29 and 41 students of mixed ages from 5 years to 16 years and mixed grades from first through eighth. Students in the experimental and control schools had similar prior education, attainment, socio-economic and ethnic background. All three districts had similar economic resources. The experimental school's enrollment exceeded the enrollment of the other two schools by 10 and 12, and so an assistant teacher was added midway in the first year of the experiment. The experimental school also had greater opportunities for community and parental involvement, along with substantial initial parental resistance to the experiment.

By the end of four years, by the standard testing measures of the time, the children in the experimental schools showed 38.1% greater gain over the children in control schools on tests of common facts and skills. Further, the children in experimental schools showed 10.8% greater gain than that expected according to national standards. The students of the experimental school

showed even more dramatic improvement than the control schools in measures of attitude toward school and education and in measures of ordinary life conduct. There were similar comparative successes in changes of parental attitudes (as measured by such criteria as visits to school, using school facilities and voting positively on education related issues) and children's participation in community life (as measured by such criteria as reading newspapers and participating in sports and agricultural clubs).

This experiment provided an overall proof of the project concept, using traditional testing measures imposed external to the curriculum. Students with the project curriculum did better than students in a traditional curriculum on tests that more directly matched the rote recitation curricula of the control schools. There were no specific links between the tests and the project curriculum, so one could make no particular judgments about how students were succeeding within the curriculum. These tests provided no particular guidance for students or instructors in the project school as to how they should proceed or how the specifics of the projects succeeded. So while this experiment did seem to support the project method, it did so entirely on the turf of traditional schooling. It developed no appropriate assessment methods to guide project based schooling and learning.

The lack of appropriate evaluation tools dogged progressive school experiments throughout the first half of the century. The Lincoln School, attached to Columbia Teacher's College, illustrates the dilemma. Although founded in 1917, as of 1934 it still did not have assessment tools that would demonstrate the particular benefits of its innovative curriculum, according to two of its curriculum leaders, although the students did well on standardized tests based on traditional curricula. (Tanner & Tanner, 1990, p. 170). In 1940, the school again reported, "We still face the problem of building evaluation instruments appropriate to our changing curriculum" (Tanner & Tanner, 1990, p. 171). The next year, unable to demonstrate its particular merit, the Lincoln School was merged with the more traditional Horace Mann School.

The Eight-Year Study, which ran from 1932 to 1940 to examine the effects of encouraging experimentation in the high school curriculum, also confronted the problems of lack of adequate assessment tools (Aikin 1942). The study identified 30 schools committed to creating innovative curricula and received assurances from major universities that applicants from these schools would be eligible for admission even though they had not completed the standard admissions requirements. With that license, the 30 schools tried a variety of educational experiments, which were then assessed (Giles, McCutchen & Zechiel 1942; Progressive 1943). The final report of this study identified success by a set of general measures for the students who continued on to college, including college GPA. They found that, overall, students from experimental schools did slightly better than those from traditional schools on all measures, and those from the most experimental schools did substantially better. But again these assessments were measuring the same terrain as those devised for traditional curricula, and all were on measures removed from *in situ* performance and activities of students within the experimental curricula.

Within the Eight-Year Study there was some interest in developing new measures that would reflect the special concerns of the schools and the study. Several new tests were developed to

assess students' development in reasoning, social sensitivity, personal and social adjustment, personal interests, and appreciation of art (Smith et al., 1942). Many of these were short-answer psychometric instruments on to identify individual characteristics rather than situated use of knowledge and learning. For more advanced scientific and logical reasoning, however, there was an attempt to measure knowledge in practice. In the first experiments students were asked to write responses to novel problem situations that could be solved using particular "principles which the teachers had selected as important for their students to apply." (p. 81) These problems were of a general sort, removed from the classroom or practical activity, but still they were designed to invoke the kind of reasoning engaged in regular subject matter learning. The papers produced, however, were not easy to evaluate in a consistent manner, in large part because so many different sources of potential error, which would indicate different reasoning difficulties. Also it was hard to distinguish what was adequate evidence. Further, it was hard to distinguish problems of expression or even problems of handwriting from problems of reasoning.

In successive iterations the test became increasingly based on short response, increasingly controlled within complex multiple-choice schemes. And the problems became increasingly abstract and atomized logical puzzles. The test became removed from actual reasoning about subject matter using subject matter content, and turned into an elaborate set of unusual puzzles with an unusual scheme of short answers. Even though these tests began to approximate standard psychometric tests, they were never successfully used to establish characteristics of students, let alone to evaluate relative accomplishment or distinguish between the value of various alternative pedagogic practices or reform models. Nor could they identify the special kind of skills learned within each academic domain. While the alternative assessment efforts are described in great detail in a separate volume, they played little role in the summary volume describing the accomplishments of schools and students in the Eight-Year Study (Aikin 1942).

With no articulated way to gain a planning and assessment link between project based activities and traditional disciplinary knowledge beyond the use of standardized tests that played by the rules of the traditional curricula, there was little glue to hold the disciplinary knowledge firmly attached to the progressive curriculum. So it regularly slipped off. That is, traditional subject matter was tested on assumptions based on traditional pedagogy, and so reinforced that pedagogy as the proper vehicle for subject matter teaching. Student-centered pedagogy pursued its own path with no particular accountability to traditional subject learning. Late in his career Dewey regularly had to speak to his own followers about the necessity of maintaining strong disciplinary knowledge within progressive education (see, for example, his late essay *Experience and Education*, 1938). Progressive education since then has regularly been tarred with the brush of absence of standards and formal learning. Progressive, child-centered education is regularly cast as the romantic enemy of high quality schools with rigorous, accountable, assessed schooling. Child-centered education is regularly seen as giving way to children's undisciplined world in sore need of the basics of disciplined knowledge that school has to offer.

Neither Dewey nor we can solve this dilemma of student and subject by insisting that teachers bring to bear disciplinary learning and by hoping that students interests will lead them into the paths of desired knowledge. While we may firmly believe that true, deep, and disciplined

learning comes only from personal engagement in serious problems, unless we can find a way to harness this idea to orderly learning and institutional imperatives for planning and accountability, it is very difficult to make this insight the basis of widespread schooling. This dilemma needs to be solved if we are to create a robust, pervasive orientation towards learners' understandings and motivations in our schools at all levels. Even if we are successful in fostering the individual talents of teachers so that they can regularly improvise ways of holding student and subject together, no institution can rely on genius every day by all participants. We need ways to identify regularizable goals, institutionally supportable activities with appropriate finite resources, strong tools for recognizing higher standards for teaching and learning, and convenient instruments for monitoring results.

## **6. The Dilemma of Assessment as Seen in the Teaching of Writing**

The teaching of writing has posed the tensions in assessment quite strikingly in the last several decades. Traditional expectations for writing were writ rather large (though not universally) in primary, secondary and university curricula at mid-century. Correctness and good form in handwriting, spelling, morphology, and syntax dominated curricula and textbooks, with some secondary interest in some formal features of paragraph and short essay organization: topic sentences, the four modes of paragraph organization, the five paragraph essay, the use of examples. Indeed in higher education students' inability to produce these formal features was frequently the criterion by which students were assigned explicit writing instruction, as assessed through either a short answer exam or a short writing sample read for correctness. Demonstrated ability to write without errors would mark the end of a student's writing obligation.

This is not to say that writing wasn't an interesting and challenging part of courses in all subject areas. Some instructors created interesting writing tasks as part of students' engagement with their subject matters. As well, some instructors provided wise guidance in how students might successfully address writing tasks. But interesting and challenging aspects of writing pedagogy were unregulated, unmonitored, never guaranteed, and not even widely discussed and supported—so that they were not generally understood as essential to learning to write. Further, writing instruction, such as it was, was hidden within the more general language arts and English curriculum of K-12 schooling and was rarely the primary focus of university courses beyond introductory levels, except for creative writing. Only by the luck of the draw would one find strong and challenging writing experiences and instruction. Insofar as university students were aware of the writing challenge of particular courses, they were at least as likely to avoid such courses as to seek the focused demand.

The lack of well-defined curricula in writing at all levels (as well as of curricula that recognize the particular challenges of writing in different disciplinary, professional, or public domains) can be attributed in part to a general perception that writing was an unsituated, universal competence. If you could spell according to convention, choose words within the dominant formal dialect, parse sentences and write parsible sentences, create topic sentences and outlines, and so on, you

could write. The variations that appeared in practice (what made one text different from another) were treated largely as matters of talent, even genius (Palmquist & Young, 1992). If, on the other hand, we understand writing to be an individual act responsive to particular situations, many aspects of writing would vary according to situation, writer's communicative impulse, writer's thought, subject, and task. Only a few basic shared elements would be regularly and recognizably present in all kinds of writing, and even these would be open to situational and rhetorical variation—such as the strategic use of unexpected word order and violations of syntactic completion. Nonetheless by studying the regularities and patterns of situations, tasks, and situationally related forms, we could give some guidance useful for students who regularly face specific kinds of writing challenges. It was, however, only a few basic aspects of writing that had been given regular attention in writing instruction, and not the skills of situational response that gave meaning and motive and challenge to writing.

Further, without writing tasks being embedded within complex interactional and intertextual situations, it was difficult to define meaningful challenging assignments within writing courses. Language challenges come from the complex demands evolved over a communal discussion over time. In courses that treat writing as an undifferentiated universal skill, assignments tend to be the beginning of a discussion rather than an intermediate point in an unfolding, motivated interaction. That is, writing assignments and tasks typically ask students to take up new topics and themes, supported by perhaps a couple of readings and some class discussion. A sequence of assignments leading up to a term paper can offer a bit more embedded and coherent experience, but still such assignments often mark only a student's first foray into an area and a debate. These "conversation starters" are very different than continuing turns in an ongoing exchange of legal briefs or advancing a line of research one has been reading on, thinking about, and working on for months or years. Conversation starters, while having their own sets of uncertainties and challenges, appropriately stay simple and unspecialized in their statements, having few assumptions and few prior issues to address. Five or 20 turns into a debate, however, issues have emerged, words have taken on particular weights; resources and assumptions have emerged; complexity, qualification, and concession become necessary; prior speakers must be respected; each participant has committed him or herself to particular stances and roles, and many other challenging complications must be attended to. Thus challenging writing assignments are both particular and embedded within domains of inquiry, deliberation, controversy, or other evolved intertextual fields, such as occur in the disciplines or professions, or as part of informed, structured controversies. For such reasons the most challenging writing assignments typically emerge within the context of disciplinary coursework. In subject-matter writing one would have to sort out issues of constitutional law, or develop theoretical implications of a sociological study, or interpret the works of a novelist in relation to the political climate and cultural marketplace of the author's time.

Assessment of writing skill was consistent with the generalized view of writing as the production of correct forms. Tests of spelling, grammatical correctness, and proper usage, often in multiple choice format were widely used for placement, diagnosis, and exit. When actual writing samples were used they typically asked for statements on subjects students had thought little about, had

few particular intellectual resources for, had not engaged in a long disciplinary discussion about, and would not continue in dialogue over. Thus student essays were framed as early turns in a conversation that they would not continue and had no substantive stake in. Further these essays were often graded for correctness, with specific points deducted for specific kinds of errors. Producing correct sentences was more an imperative than writing something interesting, important, or situationally relevant and effective. Even when graded holistically the criteria for assigning grades were based on general skills such as use of topic sentences, organization, and patterns of error and correctness. Term papers and essay exams for subject-area courses might give a bit more of an experience of an embedded writing task, but the writing aspects of these assignments were rarely placed in the foreground granted to the disciplinary subject matter of the tasks.

More interesting and challenging parts of writing were generally hidden under such vague terms as “aptness of expression,” “imagination,” “quality of thought,” “disciplinary thinking,” and “critical and analytical ability.” These mysteries were typically associated with individual talent and aptitude for particular areas of inquiry. While such talents might be spotted and nurtured by instructors of courses that contained substantive discussion of intellectual matters, they were not the appropriate topic of writing instruction and not plausible matters for standardized testing. The most one could do in a standard testing situation (which nonetheless provided whole text responses) was to recognize the unstandard essay that somehow exceeded the standard competences to be displayed. The prompts aimed at eliciting standard competences, however, would not necessarily create many opportunities for interesting answers that would display unstandard competence. Scripts identifying students with special talents might have no curricular consequences except exemption, so that students could do without writing instruction and move on to “more advanced substantive courses.” Here we have a double obscuring of advanced skill. Even when advanced skill is recognized under romantic notions of talent, little is done for the students with it. And even less is done for those identified as without talent, because talent is treated as unattainable by the ordinary student. This double obscuring keeps writing pedagogy distant from making explicable and attainable the more advanced reaches of writing skill.

With only a few basic elements of writing open for discussion and instruction, the technical vocabulary of writing assessment and teaching remained small and impoverished with respect to the full domain of choices and attention of skilled writers. The one available technical vocabulary of any sophistication that opened up any situational thought about writing was the rhetorical vocabulary, which started to gain favor as teaching of writing became more professionalized in the 70’s and 80’s. This vocabulary, while most useful to domains of argument, nonetheless, offered one way to think about writing tasks. Assessments, however, have done little to incorporate even this level of technical evaluation. While this may bespeak the difficulty of building any more advanced consideration into writing assessments where individual judgment leads to nonstandard behaviors, it also bespeaks the resistance of standardized testing of writing to reach toward any significant challenge and situated thinking.

The introduction of more advanced concepts into writing instruction and an increased level of professionalization of writing instructors at all levels lead one to hope for a raising of

writing engagement and challenge. Such improvement will only gain institutional robustness, however, when built into the schemes of requirements, expectations, assessments, and accountability that organize and hold together educational institutions. For that, then we need to develop ways of assigning, assessing, and expecting advanced accomplishment on more than an individual basis. For that we need regularizable ways of prompting performances, describing what needs to be accomplished, and technically identifying levels of performance, even as we recognize that what counts as advanced competence is variable according to circumstance, discipline or other social organizational setting, and individuality of writing choices.

Developments within writing pedagogy and theory complicate the problem of identifying advanced competence even more. Writing research, theory, and pedagogy in the last quarter century have directed attention toward the situation of the writer, toward the internal and group processes by which writing is accomplished, and toward the conditions that foster or discourage expression and communication. Such approaches have defined the dominant progressive pedagogies and understandings of writing within both K-12 and higher education. While such approaches to writing have brought human motive and meaning-making to the previously dominant approach of conventional propriety, they have turned attention away from the what is accomplished on the written page and how writers' choices may affect audiences and mediate interactions. Classical rhetoric, as mentioned previously, has offered the most widespread approach to writers' choices, but has not yet emerged into much of an institutional presence in assessment or curricular planning—beyond the presence of a few concepts like audience and persuasion into some standards and assessment rubrics.

If rhetoric were to gain general acceptance as an approach to writing it might provide some means of assessment and curricular planning based on the situated production of texts that speak strategically and persuasively to some rhetoric-appropriate tasks. Rubrics for holistic scoring can be developed around students' abilities to use rhetorical concepts to shape their texts. While this might be a large step forward, it would still not be adequate to the task at hand and might lead to mismatches between skill and assessment measures. As I have argued elsewhere (Bazerman, 2000; Bazerman & Russell, 1994) rhetoric has a limited model of the functions of writing, the social interactions it can mediate, the concepts useful to strategically shape and interpret texts, and formal textual realizations that would accomplish valued work. Rhetoric, as it has emerged historically, is directed towards high stakes agonistic public performances (primarily spoken) having to do with policy choices, the adjudication of disputes, and the forging of communal values. Such discourses are important, particularly to democratic political participation and deliberation, but the uses, forms, situations, and purposes of writing extend far beyond those that rhetoric was developed for, not least because of the complex forms of textually mediated social organization developed in the modern world. Does rhetoric give adequately full guidance for the drafting of contracts, or the design of application forms, or the production of effective history textbooks? Many of our most sophisticated, complex, and socially important forms of writing are not easily or usefully characterized in terms drawn only from rhetorical theory. Effective writers of theoretical physics articles, international economic policy analyses, information systems

planning documents, medical reports, or social and cultural criticism need to know a great deal more about writing than what rhetoric can tell them

Further, social science has given us new tools of research and theory through which to view how and what people write with what effect. Our understanding of writing can be greatly enriched if we step beyond traditional rhetorical approaches. For example, consider how the limitations of gathering information about audience response and attitudes has led to rhetoric's analysis of audience being based on the rhetor's projection of audience prior to a single moment of delivery rather than a long term analysis of the interaction of rhetor and audiences (see Blakeslee, 2000). The essays in this current collection and related previous articles have been developing an alternative way at looking at the production of writing, the form and function of the texts produced, and the mediating role texts have in socially organized activity systems. As I will elaborate in another section of this hypertext (*see [section 9](#)*), I believe these genre and activity based approaches can help us shape challenging, motivated writing curricular and tasks that provide enough reliability of expectation to develop sequence and assessment.

Sequence and assessment are extremely important in writing instruction to support writers' developing the high levels of focused skill necessary to be effective to contemporary complex literate forms of activity. As those who have become skilled in writing know, it takes many years and many challenges, and that only prepares any one writer for a limited range of genres. A lawyer who has devoted many years to writing effective criminal briefs is not necessarily an effective writer of social policy or news stories, let alone an effective writer of novels or of psychological research. Yet our tools for curriculum and assessment barely cover the earliest stages of writing, so it remains hard to distinguish higher order skills or to articulate appropriate curricula beyond doing a lot of writing and providing opportunity for peer and expert critique. In higher education the burden of identifying more advanced accomplishment most frequently occurs within advanced disciplinary courses, and then in graduate education. This may well be appropriate, for advanced discursive challenges occur only within motivated advanced activities embedded within ongoing sequences of directed interaction and tasks. Further, only disciplinary experts may have the expertise to recognize effective writing in their fields, even if they lack a technical vocabulary to identify what makes that writing effective. Only in exceptional cases do the disciplinary experts have a substantial vocabulary for dealing with writing beyond the layperson's comments on correctness. But lacking technical vocabulary, the instructor may attribute effective writing to disciplinary skill or intellectual talent apart from writing, with the consequence that the student's attention will not be directed toward improving writing and no specific writing support will be offered.

To make my point a bit more sharply, let me compare the case of advanced writing instruction to the case of advanced mathematics instruction. There are many fields that require advanced mathematical expression and calculation, just as there are many fields that require advanced writing skills. It is usually quite evident, however, which specific mathematical skills are necessary to carry on certain branches of physics, or economics, or city planning projections. No credible argument or analysis could be made in such areas unless one can handle the relevant tools. These skills are so identifiable, students are advised to take specific sequences of courses

to prepare them for the mathematical needs of particular professions. One's knowledge of each mathematical domain can be specifically assessed as well as specific needs for further mathematics education. The application of these tools to domain specific tasks as well may lead to courses like calculus for engineers. The fact that such courses are easily definable suggests how well articulated the advanced skills and their application to tasks are. The cutting edge of research or design may require developing new mathematical tools, but even this will be built on a complex defined substrate of specific forms of mathematics.

Writing curriculum to a much greater degree than the mathematics curriculum remains largely unarticulated, with an advanced course being much like a previous course with perhaps somewhat different readings and assignments. The separation of upper-division writing courses into discipline specific practices is certainly an improvement, but these are not guided by some widely shared principles that articulate the difference among what is to be taught and how in these differentiate courses. Nor is there much in the way of guideposts for writing in graduate studies, although a few courses have developed on some campuses. The widely adopted workshop model, which does provide useful opportunities for exchanging craft knowledge, does not dictate any particular analytical or conceptual tools for the improvement of writing practice. Nor does the workshop model define any particular targets for writing accomplishment or criteria for evaluation except what emerge spontaneously from the workshop group, based on their experiences.

Unless we can articulate better what we are looking for, find ways to elicit it in conditions that make these more advanced aspects of writing necessary and meaningful, and point out to students what they need to work on, writing education will find it hard to move beyond rudimentary approaches to the subject.

## **7. Alternative Assessment**

The concerns I raise here about assessment are hardly new—

- the abstraction of assessment from local meaningful activity,
- the separation of competence from situated practice,
- the atomization of knowledge within assessment instruments,
- the lack of cognitive ambitiousness and challenge of standardized tests,
- the skewing of motivation within testing situations,
- the distrust and competitiveness behind testing and accountability schemes,
- the effect of testing on the curriculum.

Many have discussed these concerns, and many have developed modes of assessment that overcome parts of these difficulties. Many of these evaluation procedures have come to be

known as performance assessment or alternative assessment. They have especially flourished in the last two decades, but the Eight-Year Study in the 30's also developed assessment tools to fit more closely with non-traditional curricula. Both that earlier and current alternative assessments have addressed testing issues of cognitive ambition, atomization of knowledge, lack of relation to actual tasks, and motivation.

Some forms of alternative assessment aim at psychometrically identifying the non-cognitive aspects of learning, such as motivation, attitude toward the material, engagement, anxiety and apprehension, and so on. Recently such tests have been of interest to those who want to evaluate individual personality so as to estimate who should be selected for various challenges. Accordingly, organizations like the Educational Testing Service that attempt to predict outcomes for individuals to improve selection for advanced education have taken an interest in such assessment. These are of great interest and utility in understanding individual student's approach to and participation in various educational tasks and the effects of various curricular and classroom arrangements on students modes of participation. Nonetheless, this form of alternative assessment still does not get at the assessment of the use of knowledge, skill, and understanding within situated activities. If we are to understand learning we have to be able to take its temperature where it is happening.

Performance assessment more directly aims at capturing student knowledge, thought, and learning in the course of accomplishing tasks. Performance assessment has come to mean testing which requires a more open-ended and longer response from the person being tested—something more than a multiple choice answer. In mathematics or mathematical sciences this might mean student-produced short answers or full sets of calculations. In other domains it might mean an extended written response. These are in fact the traditional form of examinations, class assignments, and homework. Insofar as school itself is a situated practice these long-standing assignments are as situated as they can be and as motivated as students are motivated for school. The ambition, analytic depth, standards for expression and coherence, are as great as the teachers and students can organize locally within the institutional and social setting of the school. The demands are as high and focused as the teacher can articulate and bring the students along with, and are precisely tied to the educational interaction of the classroom. The particular tools and criteria are as individualized, focused, and motivating as the teacher can make them with the students and local conditions in mind, within the degrees of flexibility allowed by the local institution. Gains to be made at this level are gains to be made directly in the classroom, in conjunction with curricular design and immediate interaction with students, as well as the flexible reshaping of assignments and activities to meet the students and stretch them from where they are in that class, in that subject, at that point in time. The *in situ* assessment of performance serves most directly to guide the work of teacher and student.

While performance assessment has always been a de facto practice in class, considering it as an organized area for study and development can lead to improved practice and tools. As an organized area of study and tool development, in fact, performance assessment has been directed largely toward formative assessment—that is, assessment done at the beginning of an educational sequence to help direct classroom activities as well as individual attention. The more

tightly these assessments are tied to ensuing work and the criteria by which that ensuing work will be judged, the more useful the assessment is for orienting both student and teacher to the task at hand.

Performance assessment has recently gained some support (Herman, 1997; Henderson & Karr-Kidwell, 1998; Meadows & Karr-Kidwell, 2001). It is now being applied more widely in evaluative assessments, including even large-scale assessment programs despite the difficulties of designing tasks, the time required for grading, the difficulties in comparability of scoring, and the overall cost. Among the benefits associated with using performance assessment is that they encourage curricula and educational programs directed towards learning based open activities rather than production of short answers for standard psychometric exams. The SAT I already includes some open non-multiple choice responses and will soon include open-response writing tasks <<http://fyi.cnn.com/2002/fyi/teachers.ednews/03/25/sat.overhaul.ap/index.html>>. The most ambitious implemented use of performance assessment has been the use of portfolios for statewide writing assessment in Kentucky since the mid-1990's.

Portfolios are an extension of performance assessment, aggregating the actual documents created across a period rather than just the numerical residue of a grade. The portfolio allows for full inspection by both teacher and student of all or selected performances, to be commented on, reflected upon by peers, students and teachers in highly individualized ways that are sensitive to and draw upon the local circumstances of production and use—that is the assignments are seen in the context of the remembered, shared experience of the class. The individuality of the student, the students' engagement with the material, learning challenges and trajectories, and the particular accomplishments are made more accessible and assessable within the rich set of artifacts collected in the folio and read within the context of the events that produced the work. Portfolios are of interest precisely because they allow a richer, more individual set of resources for individual reflection, evaluation of accomplishment, and assessment of skills. The literature on portfolios is now extensive; for example, one bibliography runs over 100 pages, and the literature continues to grow (Northwest, 1996).

Because complex performances seem to offer richer, more multi-dimensional, more cognitively ambitious and potentially more coherent and less atomized snapshots of student skills and knowledge, it would seem a good idea to use them for other assessment purposes that reached beyond the local classroom. And indeed, the teaching of writing has won over the last several decades substantial battles in using written essays for a number of assessment, entrance, placement, and exit purposes where others would wish to use cheaper, less time-consuming machine-gradable exams. In almost all major university systems where there is a choice for placement within the writing sequence, a written essay is used, and where there are exams for exiting courses, advancing to upper division, or graduating, they are almost always in the form of essays. College admissions researchers, using data from California, are also finding that writing exams (as for example the SAT II writing exam) are among the best predictors of general college success, exceeded only by high school GPA (Geiser & Studley, 2001).

However, insofar as writing assignments and other extended performances and their collections in portfolios become removed from the contexts which give them meaning and interactional value, they begin to suffer the thinning effects of other forms of decontextualized assessment. The questions have to be broad and unspecific enough to be equally intelligible and motivating for all the tested population, without advantaging any group (except those of course who have the specific skills being tested for). But these questions must do this without the context and unfolding dynamic of the interactions and relations of the classroom or other activity system. That is, these questions are conversation initiators among strangers with whom one has absolutely no stake or need to communicate beyond the evaluative function of the test. (This problem of course was one of the underlying problematics of Britton's study *The Development of Writing Abilities*, where it was found that most secondary writing in this study of British secondary education only adopted the limited author-audience relation of student to examiner. Much of composition pedagogy since then has been directed at enriching that relation either by enacting richer student-teacher relations or finding broader audiences in classroom peers or other groups outside of the class.) Further, with such broad, institutional populations to test, it is difficult to calibrate the level of ambition of the questions to elicit any but the most basic skills displays. Such decontextualized writing exams typically serve minimal placement and exit purposes only and are rarely presumed to give opportunity to display higher order accomplishment. Indeed, savvy test takers avoid more creative or ambitious responses to the tests that might not be understood by cold readers who share no context. That is, the test becomes the context and the discourse is precisely limited by the perceived aims of the test.

Further problems with decreasing the sensitivity and challenge of written tests come from grading procedures. In order to provide consistency of grades across readers and across contexts of reading, general practice now involves the creation of rubrics of ideal types to which readers would be normed through training sections. These rubrics are based on a few general characterizations, which are organized into a small number of levels, typically four to six. Typical is the 4-point scale from the California Grade 4 and 7 Writing Standards Test which is part of the STAR-9 testing program ([see California Rubrics](#)).

While these scales usually are to characterize the total quality of the essay, they can be more specialized to focus on one or a limited number of traits of interest, such as thematic organization and coherence, spelling accuracy, or sentence fluency. The California Standards Test Scoring Rubric for Grades 4 and 7 in addition to general criteria for writing includes focused criteria for narrative, summary, response to literature, and persuasive tasks. The use of these rubrics with appropriate training of the readers can lead to acceptable reliability of scores across readers (Wolcott, 1998, but see Scharon, 1996 for a critique of over-concern for reliability.) Rubrics, nonetheless, provide only a crude sorting of the essays and do not identify the particular learning or knowledge displayed in the complex competence of writing. For example, the grade 7 persuasion task criteria have only the following very general characterizations.

- authoritatively defends a position with precise and relevant evidence and convincingly addresses the readers' concerns, biases and expectations.

- generally defends a position with relevant evidence and addresses the reader's concerns, biases, and expectations.
- defends a position with little, if any evidence and may address the reader's concerns, bases, and expectations.
- fails to defend a position with any evidence and fails to address the reader's concerns, biases, and expectations.

Even more seriously, they homogenize the variety of approaches a student could take to the writing prompt, which must be characterized as fitting in one or another of the ideal types. What if a student addresses the persuasive task by framing an extended analogy that taps the common experiences of the audiences, as in the biblical parables? Or presents a compelling vision to evoke a values commitment, as in King's "I Have a Dream" speech? These would fall almost entirely outside the rubrics. Further these ideal types are characterized by formal features of the text, rather than the effectiveness of the text for the situation that prompts it. Rather the context becomes only the test-taking situation where one necessarily wants to produce the formal features identified by the rubrics. Thus exam-savvy students would necessarily adopt a claim and empirical evidence form, punctuated by concessions or counter-arguments to deal with standard points of contention, even if that would not be the most rhetorically effective strategy for the task at hand. Teaching to the exam leads instructors to teach writing as producing a formal product as described by the rubrics. This is in tension with a pedagogy that looks toward student participation in compelling rhetorical situations as the best way to elicit, practice, develop complex and rhetorically powerful choice-making. Where these rubrics define writing competence in generalized, unsituated terms, they trivialize the task of writing, rein in communicative ambition, and even remove the communicative impulse.

Portfolios of material produced *in situ* across the year within real learning tasks and then collected for an overall assessment provide one means for overcoming the artifice and lack of motivated ambition of out-of-context exams. Insofar as they remain embedded within a continuing set of educational interactions they can offer powerful tools of mutual review of development and accomplishment. But they can present similar problems of abstraction if the concern is not personal assessment and individuality but comparability and certification of a level of cognitive, disciplinary, or rhetorical competence. Although the component texts carried forward may have had strong rhetorical force in previous circumstances, they lose that context.

That context does stay in one sense insofar as the achievement of the paper was a consequence of the nature of the dialogue and expectation in the class. If the course content, presentation and discussion posed interesting questions that motivated students to want to pursue questions in disciplinary depth, and provided access and tools of access to resources, and kept raising the stakes of what would be acceptable, persuasive arguments, then the paper would reflect that knowledge. That is, the paper is an artifact of the entire activity system of the class and not just the individual merit of the student. And the writing of the paper, the learning experience, the amount the student is stretched, has to solve new problems, confront new materials, articulate

ideas with greater precision is precisely a result of this *in situ* process. It is this experience, honing of thought, knowledge, writing skill, commitment to particular ideas, sense of high level communication that the student brings forward through including the resultant paper in the portfolio. The paper, once removed from that classroom dialogue, loses the vitality (that is, the concrete life) it had there, but contains the residue markings of that dialogue. It is those residue marks of thought, learning, and skill that need to be captured if we want to know what the student learned and can now do. If the writing is truly for an audience and set of activities beyond the classroom, then the writing will gain its life within those broader circumstances. In that case the external audience can be deeply motivating. If the writing, on the other hand, finds its life only in the communicative world of the classroom, the outsider's assessor's eye may have a chilling effect on that dialogue; at the very least the local texts will need a reinvigoration to make them meaningful outside that context.

The pieces in a portfolio were once meaningful *in situ* work, but now in the portfolio students need to create a new testing context out of them. This may be done by renarrativizing—either to establish the former conditions of production or to highlight skills displayed in each. In both cases the texts become examples to be read in relation to the new narrative context. The writer might create a narrative of growth, of a line of thought, of personal awareness and self-formation or of technical skill. In this case the earlier texts are examples of lack of development rather than positive indications of skill; only the latter items exhibit learning and accomplishment. A recent dissertation by Anthony Scott at the University of Louisville has demonstrated that this renarrativizing of portfolios through cover letters directed toward the examiner has already begun in the Kentucky statewide portfolio assessment system (2002).

In most cases of evaluating portfolios, given the heterogeneity of texts and situations that generate the texts, it becomes hard to identify what we should look for in common across the texts beyond formal correctness. If we create general rubrics we can easily lose the particular character and meaning of the texts, what they significantly do in their original or even new portfolio context, and thus the principles upon which the texts are constructed. We lose the very idea and shape of the text around which all the other features are arrayed. And once again we lose sensitivity to higher level achievement, and thus remove the motive for students to display themselves at their most engaged. One way to capitalize on renarrativization is to define the criteria for the evaluation of portfolios around growth and thoughtful use of specific disciplinary knowledge, particular cognitive skills, depth of reasoning and creative/extensive use of resources. The genres and activities to be presented in the portfolio would be coordinated with those goals, and the student statements and selection of samples will be directly to create narratives of these disciplinary, higher order reasoning practices that the students have developed.

Greater comparability in portfolio samples and in the situations that produce them can also be obtained by identifying more closely the genres to be produced and the activity systems the genres these come out of. Those genres and activity systems may then be aligned with the typical genres and activity structures of schooling at the appropriate grade levels. The identification of broadly defined genres to be included within portfolios is already a practice in some cases. For

example, the Kentucky Writing Assessment specifies that at grades 4, 7, and 12, portfolios include samples of reflective writing (in the form of a letter to the reviewer), personal expressive writing (in the form of a personal narrative, memoir, or personal essay), literary writing (in the form of a short story, poem, or script) and transactive writing (to be selected from a variety of forms) (Kentucky, 2001). Perhaps we can make some gain in the situated meaning and complexity of performance tasks by seeking comparability not in the final text but in the situation and activity system that produce the text. As I will examine in the next several sections, within organized cultural practices, shared cultural understandings can lead to a set of broad expectations that will lead to similar behavior that might be comparable across situations.

## **8. The Need for a New Approach to Assessment**

Each of the previous sections of this hypertext has in its own way been articulating the need for new tools to assess motivated situated activity. Such tools could locate specific skills within more advanced and challenging performances, elicit specific more advanced performances within assessment tasks, and articulate with more challenging curricula. Such tools would help teachers plan and direct projects so that desired areas of instruction and learning will be activated in ways relevant and challenging to students within meaningful situations. Such tools would help teachers see more quickly and fully the opportunities for supporting traditional skills as they appear in emerging activities and introducing new kinds and levels of skills needed for successful completion of meaningful tasks. Finally such tools would help teachers recognize and assess students' engagement with mandated or desired areas of organized curriculum.

While I have suggested that alternative performance-based assessments have been going in the right direction, they still lack sufficient definition and regularity to serve institutional needs and to create as high a level of focused challenge as they might. The lack of a robust solution of this dilemma--serving institutional needs as well as teaching and learning needs--provides credibility for the recurring calls towards direct skills and subject matter instruction, to be monitored by standardized examinations. We see exactly those issues arise in the current climate of standards and accountability through testing.

A recent volume on assessment from the National Research Council, *Knowing What Students Know: The Science and Design of Educational Assessment*, presents the current impasse strikingly. The volume prepared by the NRC Committee on the Foundations of Assessment through a lengthy consultative process with leading educational researchers aims to review the current state of the art of educational assessment. Early in chapter 3 of this volume, four perspectives are identified as representing current thinking and research on learning. The report identifies the four perspectives as differential (individual differences that affect learning), the behaviorist, the cognitive, and the situative (which the report identifies with sociocultural theory and research). This last perspective most closely corresponds to the activity perspective taken in this essay and other essays in this collection. Concerning this perspective the report states:

Most current testing practices are not a good match with a situative perspective....From a situative perspective, there is no reason to expect that

people's performance in the abstract testing situation adequately reflects how well they would participate in organized, cumulative activities that may hold greater meaning for them. (2001, p. 64)

A one-and-a-half page discussion of the perspective is followed by a one-page discussion of the points of convergence of the four perspectives. In this section the value of the situative approach is identified by such general classroom-focused statements as

the social practices of learning emphasized by the situative approach are important aspects of education." (p. 64)

The situative perspectives can aid them[teachers] in organizing fruitful participatory activities and classroom discourses to support that learning. (p. 65)

[Both cognitive and situative perspectives] imply that assessment practices need to move beyond the focus on individual skills and discrete bits of knowledge that characterizes the earlier associative and behavioral perspectives. They must expand to encompass issues involving the organization and processing of knowledge, including participatory practices that support knowing and understanding and the embedding of knowledge in social contexts. (p. 65)

By such language the report tends to reduce the situative approach to an issue of effective classroom technique, and gives only minor nods to its more thoroughgoing claims—that knowledge and thought are mobilized and organized purposefully in situations and are not usefully or measurably assembled except in such contexts. The volume never returns to the situative approach in the rest of its 350 pages, even though it lists the situative approach as one of the four major perspectives, and is the one that is most recent. In short, the volume offers no way to deliver on the claims of situatedness for assessment, and declares it relevant only as a classroom planning concern. That position might perhaps be acceptable if assessments and assessment-based decisions did not regularly influence, shape, and even regulate the practices of the classroom. As long as the latter is the case (and there are plausible institutional reasons for that being the case—*see section 4*), the socio-cultural perspectives must find some way to find a way into the assessment system.

Not only must the sociocultural or situative perspectives find some way in, they themselves may provide the means for creating a place for themselves within the world of assessment. In the remaining sections I will be exploring the ways in which an activity systems approach to genre may provide some useful tools for assessment and planning. In line with the approach of the other essays in this collection, I am suggesting that a genre and activity based approach can help us develop regularity, anticipability, and focused challenge within motivated real tasks. These assessable tasks will be continuous with the learning work of the classroom and will not need to take time out for isolated assessment tasks, nor will they create an imperative to create test preparation that takes time away from the situated learning of the classroom. The assessed tasks will be the comparable products of the ordinary work of an organized, motivated curriculum based at the intersection of student activity and subject-area knowledge and skill.

## 9. Genre and Activity Theory

In the earlier sections of this hypertext I have been employing an activity systems approach to understand where, how, and why we carry out assessment activities and to consider the particular forms through which we carry out these assessments. In the following sections I will use the same theoretical tools of activity and genre theory to analyze a specific set of activities in one classroom over a month-long social studies unit. This analysis is aimed at showing how particular forms of knowledge and thought are associated with and displayed within particular activities and the genres produced.

In the various chapters of this collection there are many representations and explications of activity theory and its relation to genre theory. Rather than repeat material elsewhere here and in numerous other publications (see especially Bazerman, 1994; Russell, 1997a, 1997b), I will simply highlight those aspects most relevant for the problem at hand.

**Genres** are typified forms of utterances recognized as useful in circumstances recognized as being of a certain type. We coordinate our speech acts with each other by acting in typical ways, ways easily recognized as accomplishing certain acts in certain circumstances. If we find a certain kind of utterance or text seems to work well in a situation and be understood in a certain way, when we see another similar situation we are likely to say or write something similar. If we start following communicative patterns that other people are familiar with, they may recognize more easily what we are saying and trying to accomplish. Then we can anticipate better what their reactions will be if we followed these standardized, recognizable forms. These patterns are mutually reinforcing. Recognizable, self-reinforcing forms of communication emerge as genres (Miller, 1984).

As typified responses, genres also serve to further typify motives, actions, and circumstances. They become the regularized bits of systems of interactions, out of which the entire system takes on increasingly recognizable and anticipatable sequences of action. As the form of action they also become the shape our motives take on as they become realized in action.

This process of moving to standardized forms of recognizable and easily understood utterance and action is called **typification** (Schutz & Luckmann, 1973). Thus in some professions if we wish to seek a position we need to prepare a resume or curriculum vitae to list all the relevant facts and professional accomplishments of our life and to highlight our desirable qualities for the potential performer. A standard format direct us to present certain information, such as address, education and prior experience. The standard format also directs us how to present that information. Following the standard format, as well, helps the employer find and interpret the information. Further there are standard differences in format for different professions. In academic employment, publications and research take a central role, while in business listing responsibilities in each prior position and listing of specific training and skills are often important. Of course, even within the standard forms people try to express their particular characteristics and make their resume distinctive and memorable, so as to stand out from the others. Yet as soon as someone invents a new element or format

that seems to work, it is likely to be picked up by others and become fairly standard within that field. Such, for example, is the newly established practice on resumes for a number of professions of listing computer programs one is familiar with.

Genres typify many things beyond textual form. They are part of the way that humans give shape to social activity. When you are at a football game and recognize that the crowd is taking up a chant for your team, as you join in you are being drawn into the spectacle and emotions of the community athletic event. As you read and are convinced by the political pamphlet of a candidate for Congress, you are being drawn into a world of politics and citizenship. As you learn to read and use research articles of your field, you are drawn into a professional way of being and work.

Through the typification of forms, participation in social and cultural activities itself becomes typified as do the larger systems. That is, the textual forms mediate our relationship with others who are part of the activity, thereby giving regularity to our form of participation, our relations to others, and our contribution to the entire object-orientation of the activity system. Thus a letter to a newspaper editor is not only a recognizable form, but it is tied up with many other forms, such as newspaper editorials, news stories, political speeches, campaign documents, newspaper subscriptions, and many other elements of the journalistic and public spheres, out of which those spheres are constituted.

Those spheres of activity, or **activity systems**, having then been constituted, the genres then form modes of participation and motives for formulating one's participation. That is, one sees a way of participating through the letter to the editor, and one then is so moved to write the letter. While one's general communicative impulse is shaped by the generic form and motive of the letter to the editor, the specific form, content, and aims of the letter are responsive to the current situation and its placement within a history of similar recurrent situations. Further, by writing and submitting a letter we take on specific relations to the editor and author of the story we are commenting on; and if our letter becomes published we taken on an identity and relationship to the community served by the newspaper. Finally, we have carried forward the public discussion of some issue of community concern and have helped maintain the forums of public deliberation and contention.

In engaging in the communicative relations of a genre we typically need to deploy certain specifics of knowledge and thought associated with that genre. Each genre has its pattern of information displayed as well as agents and activities displayed. Bakhtin calls this informational and actional landscape of a genre its **chronotope**—or time-space (1981). Just as fairy tales are set long ago and far away in a land populated with princes who do great deeds, maidens who find themselves in distress and witches who are up to no good, so do stock reports visit accounting rooms where profits and losses are counted up and marketplaces where customers are eager for goods and services and corporations are clever at anticipating those desires. The typified informational landscape and action of every genre establishes expectations of knowledge and thought that each text in that genre must fulfill if it is to be effective in its work.

A **Genre Set** is the collection of types of texts someone in a particular role is likely to produce (Devitt, 1991). In cataloging all the genres someone in a professional role is likely to speak and write, you are identifying a large part of their work. If you identify all the forms of writing a student must engage in to study, to communicate with the teacher and classmates, and to submit for dialogue and evaluation, you have defined the competences, challenges, and opportunities for learning offered by that course.

A **Genre System** is comprised of the several genre sets of people working together in an organized way, plus the patterned relations in the production, flow, and use of these documents (Bazerman, 1994). A genre system captures the regular sequences of how one genre follows on another in the typical communication flows of a group of people. The genre set written by a teacher of a particular course might consist of a syllabus, assignment sheets, personal notes on readings, notes for giving lectures and lesson plans for other kinds of classes, exam questions, e-mail announcements to the class, replies to individual student queries and comments, comments and grades on student papers, and grade sheets at the end of the term. Students in the same course would have a somewhat different genre set: notes of what was said in lectures and class, notes on reading, clarifications on assignment sheets and syllabus, e-mail queries and comments to the professor and/or classmates, notes on library and data research for assignments, rough drafts and final copies of assignments, exam answers, letters requesting a change of grade. However, these two sets of genres are intimately related and flow in predictable sequences and time patterns. The instructor is expected to distribute the syllabi on the first day and assignment sheets throughout the term. Students then ask questions about the expectation in class or over e-mail, and then write clarifications on the assignment sheets. The assignment sheets in turn guide student work in collecting data, visiting the library, and developing their assignments. The pace of their work picks up as the assignment deadline approaches. Once assignments are handed in, the professor comments on and grades them. Similarly the instructor prepares, then delivers lectures and classes. Students are expected to take notes on readings beforehand and then on what the instructor says in class; then they study those notes on class and readings before the various quizzes and exams. Typically the instructor looks at the lectures and assigned readings in order to write questions for quizzes and exams. The students then take the exam and the teacher grades them. At the end of the term the instructor calculates by some formula the sum of all the grades to produce the content of the grade sheet, which is submitted to the registrar to enter into an institutional system of genres.

This **system of genres** is also part of the **system of activity** of the class. In defining the system of genres people engage in, you also identify a framework which organizes their work, attention, and accomplishment. In some situations spoken genres dominate, but as you move up the educational ladder and into the professional world, the system of written genres become especially important. In some activities physical aspects take on a highly visible and central role, and the spoken and written genres are peripheral or supportive rather than central. Playing basketball may be mostly about moves and ball handling, but there are rules, strategies, cheers, league organization, and newspaper reporting which engage spoken and written genres. In knowledge-based fields, such as medicine, and especially fields where the primary product is

making and distributions of symbols, such as journalism, then the activity system is centrally organized around written documents.

Considering the activity system in addition to the genre system puts a focus on what people are doing and how texts help people do it, rather than on texts as ends in themselves. In educational settings, activity puts the focus on questions such as how students build concepts and knowledge through solving problems, how instructional activities make knowledge and opportunities for learning available, how instructors support and structure learning, and how and for what purposes student abilities are assessed.

For the purposes of assessment and planning, a genre and activity system analysis identifies the character of performance situations. It identifies the typical forms performance takes within those situations, the kind of work associated with those forms and performances, and how those forms of participation are linked to other resources and participants within activities. It also identifies the kinds of knowledge expected and the kinds of uses and manipulations that information is to be put through. Genre and activity analysis allows us to see and evaluate the performance within greater contexts at the same time as it allows us to understand the contexts evoked by assigned forms and the appropriate resources, coparticipants and activity goals that would evoke robust performances. It allows us to anticipate the performance expectations of particular forms within regularized settings, and create some comparability among assessment and performance settings. In short, genre and activity systems analyses give us ways of considering the important particularities of learning and assessment situations without seeing all situations as being so particular as to be incommensurable and idiosyncratic.

The genres students are requested to write establish expectations and the level of challenge students will have to meet to realize the demand of the genre. Each genre, to be well performed, requires the use and display of specific kinds of knowledge, specific kinds of understanding and operations applied to that knowledge, and specific kinds of intellectual skills. That is, each genre carries with it implicit challenges and implicit criteria of assessment. Making explicit the challenges and specific criteria of each writing task and the genre it is to be realized in provides the means for in situ assessment integrated the larger systems of learning activity. We may not need to isolate assessment activities from the on-going work of education. In fact, understanding genres may allow us to make comparable or even repeatable assessment activities without creating stand-alone assessment activities that have no motive beyond assessment.

This integration of standardized assessment (standard in the sense of meeting particular standards and in the sense of being common across settings) with teaching and learning situations allows richer, more motivated, more multi-resourced challenges to be placed before the students. We can then assess more complex, difficult work within richer production environments.

Benefits may also flow towards planning as well as assessment, for through an analysis of genre and activity systems we can more explicitly project where sequences of assignments will take students. We can see what knowledges students will have to become more familiar with, how they will have to represent that knowledge, and what kinds of thinking students will have to do to

appropriately complete the tasks. We can see, from the holistic perspective of meaningful learning tasks, the particular disciplines of information and thought tasks and sequences of tasks will draw students into. We can provide anticipatable pathways into received organized domains of disciplinary knowledge even as we allow students to follow the logic of their inquiry. The dynamics of the genres and activity systems will direct them towards particular kinds of available resources implied within the typifications of those forms of action. Standards based planning and testing need not atomize learning environments so as to produce atomized success on atomized tests.

Finally, extended sequences of activity may be more motivating for students as well as make them more knowledgeable and practiced within the relevant domains. If we are skillful in identifying the areas and activities for the sequences of activities we can draw the students more deeply into series engagement with knowledge, thought, and knowledgeable practice. As a consequence the learning environments will be more exciting and rewarding. We will be testing students at the best, their most engaged, with their most heightened attention carrying the momentum of a sequence of motivated activity, rather than at their most disengaged skill and knowledge display elicited by isolated testing.

## **10. An Activity Analysis of Planning, Accountability, and Assessment of a Multi-disciplinary Project Based Learning Experience**

The following case demonstrates the value of considering genre, genre sets, genre systems and activity systems in evaluating the learning potential and consequences of a set of classroom activities. I would like to thank Chris Carrera for his generosity in opening up his classroom, materials, and thought to ethnographic study and to Kambiz Ebrahim for his help in collecting the ethnographic data. This example does not yet rise to a fully coherent example of local assessment following the principals I have been laying out, let alone the basis of large scale testing. However, it does make explicit the implications of the genres and activities planned and assigned by the teacher and the anticipatable uniformity of the challenges faced by the students, reflected in the orderliness of their productions. The variations of their productions are understandable and analyzable in relation to the typified expectations of the assignments that define the knowledge resources to be drawn on and the intellectual operations to be carried out.

Over a six week period during the late fall of 1998 in Mr. Carrera's sixth-grade class in a suburban California public elementary school, students engaged in a social studies learning unit on the Maya. According to state mandate (at that time and still in force) sixth year social studies should be directed toward World History and Geography: Ancient Civilizations. While the current standards ([see California Grade Six History-Social Science standards](#)) do not explicitly identify attention to the Maya as they do mandate attention to Ancient Middle Eastern, Greek, Roman, South Asian, and Chinese civilizations, they do mandate a preliminary unit on archeological studies. It is under this mandate and with this focus that the teacher developed and assigned the set of activities presented below.

In this class the unit on the Maya was to some degree integrated with simultaneous learning units in mathematics, language arts, and video production. As part of this unit they read and wrote a variety of texts. Among the texts they wrote were fill-ins on worksheets and outlines, notes on the readings, informational reports, quizzes, exams, collaboratively written video scripts about an expedition to the land of the Maya, drafts for several of these assignments, and final reflections on what they learned from the unit. These documents are the genre set of the student writing during this unit; they were collected in a file of the student work. Readings and resources included packets of information about the Maya, supplementary reference books, reference Web sites, each other's reports, and drafts of their mutually constructed projects and scripts. Students also created drawings of Mayan sports, maps of imagined Mayan cities, collaboratively built models of their imagined cities, board games about the Maya which incorporated words and text (produced in teams of two), and videos of their adventure stories (produced in teams of four or five). We can call these latter works an extended graphic genre set, though they were not collected and placed within the student work files. The fact that certain items were not included in the student files indicates a difference in the evaluation of these productions, with perhaps a difference in assumptions about the kind of work engaged in and the way knowledge is used.

In traditional terms the activity of this unit could simply be characterized as learning social studies facts and concepts with some reinforcing activities. This characterization is supported by the collection in the work files of the final reports, the worksheets, outlines, exams, and information sheets. The final reports of most students were collections of facts gleaned from handouts, textbooks, encyclopedias, and online reference materials, presented with only minimal organization and no transition between different topics and the fact sheets. Quizzes and exams equally show only the accumulation of fragmentary facts and ideas. Only a few students were able to achieve a level of articulated synthesis that gave a sense of totality of vision to their papers. On the other hand, students seemed to have understood the expectations of the genre of reports to require a collection of information. One student, Maria, in the opening sentences of her report articulates precisely this understanding of what she has to do.

Okay, before I pour all this information on you, let me introduce you to the Maya. They had six prosperous cities: Tulum, Chichen itza, Uxmal, Mayapan, Tikal, and Palenque. Got that? Great.

Here comes the rest...

They were the first people in the New World to have written records. They also had numbers. One was a dot. • Two was two dots.●●

This goes on for about 500 words presenting information on chronicles, calendars, ball games, human sacrifice, geographic and historical extent, trading, and demise. Thus almost all the papers from the class were similar in content, organization, and diction, varying mostly in length and amount of information reported.

That students had such an understanding of the task and the genre is not surprising given that the original assignment packet for this unit described this assignment only as "three page typed

describing the Maya culture.” This assignment was embedded within a much more elaborate set of activities, described below, but the specific genre of this assignment was very narrow. The narrow collection of information focus of this assignment was reinforced and supported by a number of other activities that occurred between the original assignment and the due date of the paper (December 4). First, with the assignment packet and in the days following, several handouts were distributed to the class photocopied from reference works covering history, calendar, religion, number system, sports, cities, sacrifice, geography, art and similar topics. Second, each week in class specific topics of the information were reviewed, with an informational quiz on Friday, as specified in the teacher’s planning spreadsheet (see figure A). Third, on November 9, students had to fill out a pre-printed informational outline (see figure B) on the Maya civilization concerning The Land and Region; Classic Period; Maya Knowledge. Fourth, due November 30, just before the final reports, was a research chart to be filled out by students working in pairs, first by hand on the worksheet, and then transcribed on a spreadsheet. For five cities, each pair of students had to identify the location, record an important discovery, describe the region, and select an interesting cultural fact. Maria and Sau-lin completed just such a research chart (see Figure C).

Figure A

**ROOM ONE  
WEEKLY ASSIGNMENT  
(11/16-11/20)**

Good morning Wonder People,  
I hope you used this weekend to work on your Ancient Maya project. Today we will review your progress and begin to organize our Maya simulation activity. By the way, your ancient city models are some of the most creative works anyone could ask for.  
Since this is an early day week because of conferences, we will not have PE unless we are caught up on other assignments. Let’s stay on top of things and have another happy productive week.

	NEWS 8:30-9:00	SSL 9:00-9:30	LANG ARTS 9:30-10:50	SOC SCI 11:10-12:00	MATH 12:00-12:35	HOMEWORK all day & night	Notes
Mon.	Today we will look at three stories from the newspaper.	new vocabulary and phrases,	journal writing, reader, creating with our imagination	Ancient Maya Culture: Classic Period. +Maya Cities	math drill, problem solving, measurement, Maya math	spelling, math: due Friday, SSL packet. Maya Project	
Tues.	you will use the food section from the paper for SSL	vocabulary phrases translations	read journal stories: edit and revise.	Maya: the people, the culture, the art: Classic Period	math drill, problem solving, measurement, Maya math	check your homework assignments	
Wed.	creating a news broadcast	review the food section from the paper for SSL					
Thurs.	art lesson w/ Ms. Clevenger	art lesson	illustrating your words	a videotape production: script, setting, props	math drills, (diaries)	packets due tomorrow. Read before going to sleep. Good night	
Fri.	week in review. news as journalism	weekly Spanish quiz, Simon dice, "..."	spelling test, review weeks work	Maya: the people, the culture, the art. QUIZ!!!	math quiz, activity time, check folder (diaries)	watch the news, read, & have fun with your friends and family	

# Ancient Maya Outline November 9, 1998



You are to complete this outline with information from our Maya packet, classroom discussions, and research materials. Remember to keep the information brief and to the point.

## I. Maya Civilization

### A. The Land and Region

1. Harsh living conditions
2. Jungle, rain forest
3. Mountains
4. Mexican

southeastern states, Yucatan Peninsula  
Guatemala, south into Guatemala, northern Honduras

### B. Classic Period

1. Beginnings of Mayan greatness 300 AD
2. Flourished until 900 AD
3. Schools
4. Markets for trading goods

centers for practicing religion.

### C. Maya Knowledge

1. Master astronomers
2. " " mathematicians
3. " " architecture
4. " " writing
- " " time & calendars

Figure B

city name	city location	research information	region description	interesting facts
Uxmal	Northwestern part of the Mexican Yucatan.	The magnificent architecture here is adorned with many elaborate decorations and bright colors.	rugged terrain, and hot ground	The Magican's Pyramid has said to be built in one night.
Tikal	The middle of dense jungle, north of Guatemala.	This sprawling city consists of numerous residences, temples, pyramids, and ball courts.	viny, and very colorful	Played a soccer type game called Pok- a- tok.
Tulum	The coast of the Caribbean Sea.	Tulum prospers because it can acquire trade goods from the sea.	Water, coast, and very colorful.	Sacrificed humans to Gods
Chichen Itza	North-central part of the Mexican Yucatan.	Chichen Itza has grown to great wealth and power because of it's central among the Maya trade routs.	North-central rain forest.	Attempted to kill other cities
Palenque	East part of Mexico	Human sacrifices were located here.	Thick Mexican Jungle	They had special Ceremonies.

Figure C

The products represented in both figures *B* and *C* are mechanically organized sets of factual fragments, selected and transcribed from the distributed informational sheets. The further transcription of this material onto a spreadsheet practices technical skills, but also reinforces the idea that information (and research) consists of such fragments organized into formal categories. Thus it is not surprising that students understand the final research report as a collection of loosely organized facts. Such a task can practice useful grade appropriate skills of identifying and organizing information, but does not engage higher levels of synthesis, analysis, or discussion.

The apparently student-produced genres of outlines, worksheets, and quizzes are in fact collaboratively produced with the teacher in the specific sense that the words on the final page include words of both the teacher and the students. The teacher produces the topics, categories, and structure for the outline and chart and the questions for the quiz. He further produces the instructions on each of the assignment sheets. Also in the teacher's structuring of the intermediate informational assignments we can see the teacher's hand in the final reports. Thus these genres are strongly shaped by the teacher's decisions of what should be written and how. The students' recognition of the teacher's speech act of assignment shapes their further actions in fulfillment of the assignment, just as the teachers' further assignments are dependent on his recognition of the students' completion of prior acts. Each new student production is dependent on them having completed earlier acts, turning them into accomplished social facts that they could then rely on and build upon.

In two collaboratively produced teacher student genres, however, the teacher's decisions structure a very different kind of work for the students. First is the unit final exam, given on December 11, with three questions.

1. What qualities do you think gave strength to the Mayan Empire?
2. In what ways can trade between cities help to create good relationships?
3. Why do you think the Mayan Empire did not go on forever?

Each of the three questions requires students to think evaluatively, causally, and critically, and most of them did so. Maria provided one of the more elaborated set of answers, but not all that different from that of most of her classmates. In answer to question 1 she wrote,

I think that the accuracy in their calendars, their knowledge of the movements of the stars, their ability to create their own letters gave strength to the Mayan Empire. I also think that no matter what role you had, or what you did, you were important to the Mayan Empire, and that gave strength to the Mayan Empire.

How did such questions and answers count as an appropriate test of what the students had learned if the earlier activities were primarily transcription of fragmented information? And where did the students get the ideas and stance from which they could answer these questions?

Before we answer that let's examine another end of the sequence document, the "Final Thoughts" worksheet filled out two days before the final exam. The following example from Desmond covers typical themes.

Room One

Ancient Maya Civilization

### **Final Thoughts**

Think about all that we did with this study of the Ancient Maya: the research report, art projects, model making, script writing for the plays, videos, videotaping, and group organization. Now share some of your final thoughts by responding to the questions below. Please be Specific. Thanks for doing a great job with your assignments.

1. What did you learn from our study?

I learned that the Maya were very Bright people because they had writing, langue, and calenders.

2. What did you like about our study?

I liked making are clay mayan citys because I had a fun time working on in with my friends and I.

3. What would you like to change with what we did?

I would like to change the Mayan city

time to work on it, I would want more time to work on it. I think it would have been better if we had more time to work on it. But it still turned out good.

4. What would have made it better and more interesting?

Are play would have been better if we had more cooperation.

5. How could we have improved our video productions?

It would have been better if we were organzed.

Include some of your personal thoughts below that may not be asked in the five questions.

It was very fun.

Only the first question really evokes in Desmond (and most of the other students) any reference to the factual information, and even then the information is subordinated to an evaluative conclusion. All his remaining responses (as did the responses of most of his classmates) referred to the other activities of building models, the play production and videotaping. And key themes were working together, doing things better, and having fun--all issues of participation and engagement. Given the predominant flavor of the work we have examined so far, how did students glean such learning and develop such attitudes toward the unit?

The answers on the final exam and “Final Thoughts” reflect some class discussion about the factual material they were learning, but they also reflect the wider system of activity built into the unit. The unit was built around two sets of activities organized by the teacher—one individual, informational, and reporting; the other creative, inquiry-based, and collaborative. Each set had its own supportive and assigned genres that developed and rehearsed the assignments’ orientation, creativity, and thought. The informational content was embedded within these activities that engaged the students and which they found fun. But even more these activities gave students the opportunity to think about and use the factual content, and thus to develop significant meanings from the content.

The activities were set in motion by the original assignment sheet at the beginning unit, which set out the following simulation frame:

**Project:** You are a member of an ancient Maya people and you have been assigned the task of establishing a new site to design and build a great city.

The name of the city will be chosen from one of the following: Tulum, Chichen Itza, Uxmal, Mayapan, Tikal, or Palenque. The task is to be done individually, but you may confer with others to get ideas or give suggestions. Good luck and begin immediately, because the king is not a patient man and needs the city built before invaders arrive.

The sheet goes on to specify three parts of the project: a “three page typed report on Mayan Culture, an illustration/graphic, and a blueprint of the Mayan City with everything labeled.” A fourth final activity of group creation of a play with script and costumes and videotaping is mentioned. Each of these four parts was modified and elaborated in the ensuing six weeks.

The original situation frame of designing a new Mayan city gave motive and purpose to the informational and other activities of the first half of the unit. The factual information is necessary to understand what a Mayan city is and how you should design one to include its typical buildings, institutions, and places for its usual activities. That work became most fully and directly expressed in the map/design each produced, which then became the basis for a scale model. Two additional art projects, however, reflected the same kind of civilization building thinking. One was a board game each had to design to reflect the daily life of residents of the city and the other was to act as the chief Maya artist commissioned to create a design that reflects the style of the culture. (Students were also learning to use graphic software as part of this assignment.) Finally there was a sequence of Mayan math exercises (from a prepared unit) that used standard word and logic problems incorporating objects and situations relevant to the Mayan agriculture, social structure, and culture, the problems also provided practice using the Mayan number system and calendar. These immersions in Mayan life through simulations did more than rehearse some factual material about the Maya, they drew sixth graders into thinking about the material and how the facts reflected a way of life. Such thinking could be displayed and assessed in response to the questions asked on the final exam and “final thoughts.” Such thinking also provided the background for another level of activities in the second half of the unit.

The second half of the unit transformed the situational frame from design into inquiry and the mode of work from individual and collaborative. This was initiated by an assignment sheet handed out four weeks into the unit on November 20, just after the designs and scale models were finished. The assignment sheet informed the students that they were archeologists who had found an artifact with a map to an undiscovered Mayan City. They were to organize in teams to search for the city and its treasures; they would then script and produce a video documentary of their adventures. The assignment sheet provided space for the students to sketch out preliminary ideas about setting, characters, events, and story summary for the initial work sessions with the collaborative group (about five students in each group). Also provided was a follow-up framework for the script, in which the characters, setting for each scene, the props and costumes and the production team roles, and other notes were to be listed. These assignment sheets scaffolded the work of script writing and production for the students as they made decisions in filling out the blanks and then did the additional work implied in each of their answers.

The research chart discussed earlier finds its meaning within this archeological frame of action. The instructions for the chart describe it as a report from field archeologists back to their colleagues to let them know what has been found. So now the material is not just information to be tested on--it is something the students, in their simulated roles as archeologists, know to be shared with others. The knowledge they have found also becomes subject and material of their videos (which were also produced as live plays).

The scripts of the videotaped performances are pretty basic, involving archeologists walking through the city with local informants pointing out aspects of the culture. The plays tended to dwell on the ball game with a death penalty for losing and other moments of human sacrifice. Nonetheless, stories are larded with the facts and names that have cropped up in the various reading and writing genres throughout the unit, so that the students have learned to inhabit the informational space even while engaged in imaginative play. Looking at the limitations of the scripts, one could well understand why a number of students commented that the videos would have been much better if they had learned to work together and everyone learned to do their part. It also becomes evident that the teacher used the lesson of cooperation within successful civilizations to help students reflect on the difficulties of their own collaboration--and thus comments about cooperation being essential to Mayan success turn up as well on the final exam.

When we look at the total activity system of the classroom as students participated in each unit, and the kind of work and learning accomplished in the production of each of the teacher-directed genres, we can see that students were doing more than reproducing facts from handouts and books. They were thinking about the material and using the material to engage in other activities, which required understanding and elicited motivated engagement. These various activities were coordinated in a mutually supported sequential system that ended with classroom presentation of reports, airing of the videos produced by each of the several small groups, reflective observations on the activity, and analytical thought on the final exam. The activities each were centrally engaged with well-known, typified textual and graphic genres that afforded students anticipatable access to information, challenges and problem solving. Each activity also provided structured opportunities for learning. The end result included familiarity with some factual information about the Maya. The result, however, also included a sense of what Mayan life was like, an experience of being an inquirer into another culture, increased skill in synthesizing and presenting information, using knowledge creatively for imaginative productions, and a sense of the practical import of the information. There were also learning and practice of many computing and video media skills. The richness of the activities and the informational base drawn on combines with the regularities of the genres assigned to create engaging, informed work which is both plannable and assessable. The teacher through careful construction of activities and assignments could anticipate what kinds of intellectual challenges and knowledge resources students would address at each juncture. Further, each student production could be evaluated on the basis of how well it met the informational demands of the genre. While the assessment of the more complex collaborative assignments of the second half was more holistic than the examination and grading of the more factual first half material, the

kind of genre and activity analysis developed here provides means of developing more explicit criteria of assessment and specific guidelines for project development.

## **11. Discussion of the Case Study**

What we see in this complex of classroom activities is of course an admirably well planned set of tasks orchestrated by a skilled teacher, engaging students in attractive projects, but also providing them the intellectual and factual tools to accomplish those tasks. The students' need for the information and engaged interest in the projects seem to have motivated greater familiarity and attention to the traditional materials. The traditional materials were assessable through standard information-recall testing and other knowledge display assignments such as the research paper. However, the motivated engagement of the project activities provided for a deeper understanding of the material, as students used factual material in their roles as designers and archeologists. This deeper understanding was to some degree assessed through questions calling forth such thought and was to some degree assessed through the holistic grading of the collaborative projects, but more explicit criteria and assessment practices can be developed through further description and understanding of the tasks students are asked to perform.

The genres of activities and display typically call forth the kinds of knowledge and skills we see practiced and displayed. It is the judicious selection, combination, and sequence that supports students in building knowledge, developing cognitive complexity, and becoming engaged in learning. The timing and selection of products and assessments--traditional and non-traditional--allow students to demonstrate knowledge of these materials, and allow the teacher to build on these skills in further lessons. Subsequently, Mr. Carrera introduced even more complex and advanced units on other ancient civilizations. The state-mandated curricular survey of cultural heritages is organized as parallel monthly tours of different civilizations, differing only in the factual particulars and historical themes, to be approached at the same level. However, because of the attention to deepening activities across the units, building on the skill sets, knowledge, and genred understandings presented in each unit, Mr. Carrera was able to lead students to more sophisticated activities and understandings over the course of the year. Students demonstrably were developing a deeper understanding of culture and society as they had to accomplish more and more complex tasks in each sequence of activities and sets of assigned genred performances.

The judicious selection of activities, artifacts, projects, resources, and media also allowed Mr. Carrera to create a highly integrated curriculum drawing together several traditional curriculum areas (language arts, mathematics and social studies) with new and highly motivating areas of media skills and media criticism. In so doing he started to develop in the students an understanding of the role of representation and knowledge making on the parts of cultural recorders such as archeologists, historians, film-makers, and even students writing fact-collecting research papers.

This curriculum was the result of many years of experience and development by Mr. Carrera. He has drawn on numerous materials and projects developed by other teachers, has located and used special resources--such as the educational support services of the local cable provider--and he

has developed a keen and complex way of integrating assignments. I do not claim that he needed anything like an explicit knowledge of genre and activity systems to do this. Yet he clearly has a practical knowledge of these things because he was able to manage them so well.

An activity and genre analysis helps unpack and describe his curricular plan and sequence of events. The analysis in characterizing the genres the students worked in and how they deployed factual material to accomplish those genres also identifies exactly what students are doing and what knowledge and thought they are displaying. It also identifies the appropriate implicit criteria by which each assignment might be assessed, given the logic of the genre and activity—that is, the identification of task appropriate skills and resources deployed in the completion of each of the genres suggests specifically what might appropriately be evaluated in each of the tasks. The genre and activity analysis provides a way of seeing the disciplinary and intellectual requirements of each task and suggests how assessment can be ordered and regularized in ways accountable to disciplinary knowledge and canons of intellectual accomplishment. Finally the analysis provides means for describing and translating the learning unit and its learning goals to other classrooms. Making the learning logic of the assignments and their sequence accessible, suggests the possibility of coordinating the work of different classrooms, and comparison of the accomplishments of students in these classrooms, although that goes beyond the scope of what was observed in this classroom. This comparability would depend on the degree that any instructional unit, developed within a particular situation, may be successfully implemented in multiple sites.

As a researcher and member of a support collaborative, I at first had a hard time coming to understand all that Mr. Carrera and his students were doing. When beginning this field study, all I saw was a number of remarkable products. And even after numerous meetings and discussions, I was unable to get more than fragmentary pictures of what was going on. The curriculum seemed very much a personal accomplishment of the teacher, not through any secretive desire on the instructor's part but because none of the more familiar curriculum description tools allowed the drawing of the coordinated picture.

Only as I started to lay out the genre and activity analysis contained in the previous section did I gain an understanding of all that went into the curricular plan, the kinds of activities carried out by the students, and the kind of learning developed and displayed in each of the products and performances. Genre and activity analysis provides tools to unpack the complexities of situated learning, to plan and organize activities that incorporate spontaneous engagement with difficult challenges and rich resources. Yet these complex, engaged activities still produce anticipatable outcomes that can be assessed in an orderly way, that supports comparability across situations.

## **12. Importance of Issue and Implications of This Line of Solution**

The stakes in assessment go far beyond best way of testing the knowledges of individuals. But even accurate individual assessment requires eliciting performance at the highest level of

engagement and accomplishment a student can produce. We would not measure the artistry of a Pavarotti or a Sting by asking them to sing an unadorned rendition of “Happy Birthday” in chorus. While some of the character of their voices might still sneak through, there would be much less to distinguish them from basically competent singers. Nor would they produce peak performances stretching their limits, challenged by interesting or difficult tasks. So why would we think that denuded basic tasks would give us a reasonable measure of what students can accomplish?

However, there are even more significant stakes—for what people are challenged with determines what they will be taught and what they will learn. If there were no Verdi arias or large popular music markets, would either singer have worked on their skills and pressed their creativity as they did, in the directions that they did? To address the demanding and intricate scenes of real performance and meaningful assessment for both, they have developed their talents in extraordinary ways. The reality of audience, the responsiveness to audience pleasure, the desire to make something beautiful for audience appreciation, the complex resources of the musical traditions being brought forward—these make musicians exceed themselves by learning their craft and putting together creative, exciting performances. Our scenes of assessment are what shape the learning. We all have heard of teaching to the test: Where the test is weak the teaching is weak, so it is important to have valuable tests to teach to. But there is an equally important phenomenon of learning to the test, to the scene of assessment. If the scene of assessment engages attention and calls on complex resources to produce exciting performances, students grow. If the test is humdrum, not only do teachers drum in the old hum, the students learn to hum along with barely half a mind.

On the other hand if life gets too interesting, how can motivated, creative performances be compared? Well, it is easier to compare the excellences of a Pavarotti to a Domingo than to an Elvis. Pavarotti and Domingo have learned to sing the same kinds of songs, based on the same discipline of voice training appropriate to their form of art. They have sung similar roles in similar operas in similar halls with similar audiences and audience expectations. Their performance worlds are very far from that of Elvis. This would be also in assessing aspirants to each of their arts. The young opera singer measures herself against her models and heroes and the young rocker to an entirely different set. The genres they sing in and the nature of their activity provides guidance for comparability, assessment, modeling and aspiration. And the genres of performance provide guidelines for the training, practice, and planning of performance for each performer.

What can get tested in a reliable, specifiable way defines what people will teach. But what gets taught also depends on what is plannable. No matter how motivating, meaningful and mind expanding some experiences are, unless they can be anticipatably planned, they will remain unplanned, awaiting the spontaneous moment. And the spontaneous moment may find it hard to claim space within the planned curriculum. Unless we find ways to make interesting, meaningful, growth-inducing activity plannable, they will regularly lose out to the easily planned presentation and rehearsal of detached skill and information. Activity and genre theory offer tools for anticipating the learning consequences of offering particular tasks for student

engagement, for structuring assignments, for bringing together various elements within a classroom--books, lectures, group work. They allow us to view the classroom as a communicative system within which students face particular challenges with the support of structured resources to accomplish specifiable complex achievements. That structured space of challenge and achievement defines the zone of activity and development the students have the possibility of being engaged in. Such tools increase the likelihood that students will gain a comprehensive, disciplined, practiced skill in the traditional knowledges and arts of our society at the same time as they understand the value of those arts and knowledges for carrying out their own, motivating, personally meaningful activities.

A genre and activity analysis can also enable us to create more effective complex performance evaluations that are comparable across situations. These assessments can then also coordinate in more interesting ways with state-wide and other jurisdiction-wide standards, frameworks, and curricula. Industrial arts teachers have some understanding that to make a table specific operations will have to be worked on, certain skills employed, specific tools will have to be mastered, and even specific forms of cooperation will have to be worked out on a project team. In disciplinary subject matters we can perhaps also gain as realistic a sense of how projects can follow on each other and what challenges new projects will create for students. We may not need to leave more advanced forms of learning to the random fate of totally emergent projects or only to the art and experience of the teacher (as useful as that is) in responding to the opportunities of the moment, as useful as that may be. An activity and genre analysis can help us articulate the structures by which a skilled teacher creates clusters of activities that foster specifiable forms of learning. At the very least, this kind of analysis helps people think through what needs to be on hand to help students through the anticipatable courses of their activities. And then, at the end, activity and genre analysis will provide us a better and more specific view of what they have in fact done.

## References

- Aikin, W. M. (1942). *The story of the eight-year study. Adventure in American education* (Vol. 1). New York: Harper & Brothers.
- Allison, Libby, Bryant, Lisbeth & Hourigan, M. (Eds.) (1997). *Grading in the post-process classroom*. Portsmouth NH: Boynton/Cook Heinemann.
- Anson, C. (Ed.) 1989. *Writing and response*. Urbana IL: NCTE.
- Bakhtin, M. (1981). *The dialogic imagination*. Austin TX: University of Texas Press.
- Bazerman, C. (1994). Systems of genre and the enactment of social intentions. In A. Freedman and P. Medway. (Eds.), *Rethinking genre* (pp. 79-101). London: Taylor & Francis.
- Bazerman, C. (2000). A rhetoric for literate society: The tension between expanding practices and restricted theories. In M. Goggin (Ed.), *Inventing a discipline* (pp. 5-28). Urbana IL: NCTE.

- Bazerman, C. & Russell, D. (Eds.) (1994). *Landmark essays in writing across the curriculum*. Davis, CA: Hermagoras Press.
- Becker, H. S., Geer, B. & Hughes, E.C. (1968). *Making the grade: The academic side of college life*. New York: Wiley
- Blakeslee, A.M. (2000). *Interacting with audiences*. Mahwah NJ: Erlbaum.
- Brecht, B., (1964). *Baal, A man's a man, and the elephant calf: Early plays*. New York: Grove Press.
- Britton, J., et al. (1975). *The development of writing abilities (11-18)*. London: Macmillan.
- Bruner, J. (1990). *Acts of meaning*. Cambridge: Harvard University Press.
- California Department of Education Grade Six History-Social Science Content Standard. Retrieved July 20, 2002, from <http://www.cde.ca.gov/standards/history/grade6.html>
- California Department of Education Grade California Writing Standards Test: Grade 4 and 7 Writing Rubric. Retrieved July 20, 2002, from <http://www.cde.ca.gov/statetests/star/resources.html>
- Chamberlin, D., Chamberlin, E.S., Drought, N.E. and Scott, W.E. (1942). Did they succeed in college? *Adventure in American Education (Vol. 4)*. New York: Harper & Brothers.
- Clark, H.J. (1992). *Arenas of language use*. Chicago: University of Chicago Press.
- Coe, G. (1924). *Law and freedom in the school*. Chicago: University of Chicago Press.
- Collings, E. (1924). *An experiment with a project curriculum*. New York: Macmillan Co.
- Collings, E. (1931). *Progressive teaching in secondary schools*. Indianapolis: Bobbs-Merrill.
- Cooper, C. & Odell, L. (1977). *Evaluating writing: Describing, measuring, judging*. Buffalo NY: NCTE.
- Devitt, A. (1991). Intertextuality in tax accounting: Generic, referential, and functional. In C. Bazerman & J. Paradis (Eds.), *Textual dynamics of the professions* (pp. 336-380). Madison WI: University of Wisconsin Press.
- Dewey, J. (1900). *The school and society*. Chicago: University of Chicago Press.
- Dewey, J. (1902). *The child and the curriculum*. Chicago: University of Chicago Press.
- Dewey, J. (1938). *Experience and education*. New York: Macmillan.
- Erickson, H.L. (1995). *Stirring the head, heart, and soul*. Thousand Oaks, CA: Corwin.
- Fulwiler, T. (1987). *Teaching with writing*. Upper Montclair, N.J.: Boynton/Cook.

- Geiser, S. & Studley, R. (2001). *UC and the SAT*. University of California Office of the President.
- Giddens, A. (1984). *The constitution of society*. Berkeley: University of California Press.
- Giles, H.H., McCutchen, S.P. & Zechiel, A.N. (1942). *Exploring the curriculum. Adventure in American education* (Vol. 2). New York: Harper & Brothers.
- Graves, D. (1978). *Balance the basics: Let them write*. New York: Ford Foundation.
- Henderson, P. & Karr-Kidwell, P.J. (1998). *Authentic assessment*. (ERIC Document Reproduction Service No. 418140)
- Herman, J.L. (1997). *Large-scale assessment in support of school reform: Lessons in the search for alternative measures*. Center for the Study of Evaluation Standards and Student Testing. (ERIC Document Reproduction Service No. 435631)
- Hillocks, G. (1995). *Teaching writing as reflective practice*. New York: Teacher's College Press.
- Katz, L. & Chard, S. (2000). *Engaging children's minds: The project approach*. (2<sup>nd</sup> ed.) Stamford CT: Ablex.
- Kentucky Department of Education. (n.d.) Writing Portfolio Assessment. Retrieved December 20, 2001, from <http://www.kde.state.ky.us/capd/curric/portfolios/prtrq99.html>
- Kilpatrick, W.H. (1951). *Philosophy of education*. New York: Macmillan Co.
- Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin Co.
- Lawson, B., Sterr, S. & Winterowd, R. (Eds.). (1990). *Encountering student texts*. Urbana IL: NCTE.
- Meadows, S. & Karr-Kidwell, P.J. (2001) *The role of standardized tests as a means of assessment of young children*. (ERIC Document Reproduction Service No. 456-134)
- Macrorie, K. (1968). *Writing to be read*. New York: Hayden Book Co.
- Macrorie, K. (1970). *Telling writing*. New York: Hayden Book Co.
- McMurry, C.A. (1920) *Teaching by projects: A basis for purposeful study*. New York: Macmillan.
- Miller, Carolyn (1984). Genre as social action. *Quarterly Journal of Speech* 70, 151-167.
- Murray, Donald M. (1968). *A writer teaches writing: A practical method of teaching composition*. Boston: Houghton Mifflin.

- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington DC: National Academy Press.
- Northwest Regional Educational Lab 1996 bibliography of assessment Alternatives: Portfolios. Portland (ERIC Document Reproduction Service No. 409371)
- Palmquist, M. & Young, R. (1992). The notion of giftedness and student expectations about writing. *Written Communication* 9:1, 137-168.
- Passe, J. (1996). When students choose content: A guide to increasing motivation, autonomy, and achievement. Thousand Oaks, CA: Corwin Press.
- Progressive Education Association. (1943). *Thirty Schools Tell Their Story. Adventure in American Education* (Vol. 5). New York: Harper & Brothers.
- Rance, P. (1968). *Teaching by topics*. London: Ward Lock Educational
- Ravitch, D. (Ed). (1995). *Debating the future of American education: Do we need national standards and assessments?* Washington DC: Brookings Institute.
- Rohman, D. G. & Wlecke, A.O. (1964). *Pre-writing the construction and application of models for concept formation in writing*. East Lansing MI: Michigan State University.
- Russell, D.R. (1997a). Rethinking genre in school and society: An activity theory analysis. *Written Communication*, 14:1, 504-554.
- Russell, D.R. (1997b). Writing and Genre in Higher Education and Workplaces. *Mind Culture and Activity*, 4:4, 224-237.
- Russell, D.R. (1993). Vygotsky, Dewey, and externalism: Beyond the student/ discipline dichotomy. *Journal of Advanced Composition*, 13.1, 173-198.
- Scott, T.D. (2002). *Dear Reviewer: Reflective Writing as an Institutional Genre*. Doctoral dissertation, University of Louisville.
- Scharton, M. (1996). The politics of validity in assessment of writing: Politics, policies, practices. In E. White, W. Lutz and S. Kamusikiri (Eds.), *Assessment of Writing: Politics, Policies, Practices* (pp. 53-75). New York: MLA.
- Schön, D.A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schutz, A. & Luckmann, T. (1973). *The structures of the life-world*. Evanston IL: Northwestern University Press.
- Smith, E. R., Tyler, R. et al. (1942). *Appraising and recording student progress. Adventure in American education* (Vol. 3). New York: Harper & Brothers.

- Stotsky, S. (Ed.). (2000). *What's at stake in the K-12 standards wars: A primer for educational policy makers*. New York: P. Lang.
- Straub, R. (2000). *The practice of response: Strategies for commenting on student writing*. Cresskill, NJ: Hampton.
- Tanner, D. & Tanner, L. (1990). *History of the school curriculum*. New York: Macmillan.
- Tanner, L. (1997). *Dewey's laboratory school*. New York: Teacher's College Press.
- Tenenbaum, S. (1951). *William Heard Kilpatrick, trail blazer in education*. New York: Harper & Brothers.
- Vygotsky, L. (1978). *Mind in society*. Cambridge: Harvard University Press.
- Vygotsky, L. (1986). *Thought and language*. Cambridge: MIT Press.
- Webster, E.H. & Smith, D.V. (1927). *Teaching English in the junior high school*. Yonkers, NY: World Book Company.
- White, E., Lutz, W. and Kamusikiri, S. (1996). *Assessment of writing: Politics, policies, practices*. New York: MLA.
- Wolcott, W. (1998). *An overview of writing assessment*. Urbana IL: NCTE.
- Zak, F. & Weaver, C. (Eds.). (1998). *The theory and practice of grading writing*. Albany: SUNY