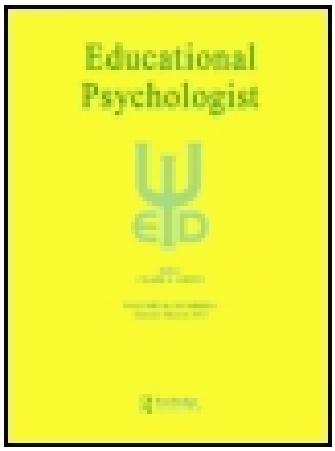

Educational Psychologist
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hedp20

Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments

Carol Kehr Tittle

Published online: 08 Jun 2010.

To cite this article: Carol Kehr Tittle (1994) Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments, Educational Psychologist, 29:3, 149-162, DOI: 10.1207/s15326985ep2903_4

To link to this article: http://dx.doi.org/10.1207/s15326985ep2903_4

EDUCATIONAL PSYCHOLOGIST, 1994, 29(3), 149–162

Toward an Educational Psychology of Assessment for Teaching and Learning: Theories, Contexts, and Validation Arguments

Carol Kehr Tittle
Graduate School and University Center
City University of New York

A framework for an educational psychology of assessment for teaching and learning is proposed, consisting of three dimensions: epistemology and theories, the interpreter and user, and assessment characteristics. The dimension of interpreter and user is equal in importance to theory and assessments, responsive to cognitive constructivism and the construction of meanings and beliefs, as held by teachers and students in practice contexts. Illustrations of the lines of inquiry and evidence that follow from this framework are given, drawing on research with teachers and using a particular assessment. Validation arguments for assessments in a practice-based context will be stronger when they are proactive and include evidence on the constructions of teachers and students and the meanings and use an assessment has for them in their educational situations.

In this article I propose a framework for an educational psychology of assessment for teaching and learning. This framework takes into consideration the current directions of the field, identifying boundaries and also supporting expansion of the set of theories relevant for an educational psychology of assessment. The importance of contexts is also recognized, and the implications of theories and contexts for the validation of assessments is developed.

CURRENT DIRECTIONS OF THE FIELD

There is an ongoing dialogue about the definition of educational psychology as an academic discipline (Scheurman, Heeringa, Rocklin, & Lohman, 1993). Two general perspectives can be identified, with a third perspective emerging. One perspective defines educational psychology as the psychology of education and the scientific study of psychology in education (e.g., Lesgold & Glaser, 1989). A second perspective defines educational psychology as an applied field that takes on the task of testing the meaning and use of psychological theories or models in educational practice (Klausmeier, 1988).

A third, emerging perspective is important to defining an educational psychology of assessment for teaching and learning and to validation arguments. This emerging perspective is interdisciplinary, practice oriented, and field based. Several factors encourage this third perspective. These factors include renewed attention in psychology to the nature of education and schooling (Berliner, 1992), the increased prominence of other disciplines in the study of education (Shweder & Sullivan, 1993), and trends in the professional, curricular, and discipline organizations related to schooling and its reform (e.g., National Council of Teachers of Mathematics [NCTM], 1989, 1991) and national efforts in setting goals and standards.

The existing domains in psychology are fragmented—learning and cognition, development, instructional psychology, motivation, personality, social psychology, individual differences, and research methods/measurement. Thus, educational psychologists interested in practice have needed to interpret, integrate, and transform research for the practice of education. The extent of the transformations required are evident in the research literature on teaching and learning in the subject matter areas, as found in the third edition of the *Handbook of Research on Teaching* (Wittrock, 1986). The need for integration is evident in the research on urban schools and their diverse students (cf. the comments of Carroll, 1993, on educational psychology and the national goals for the year 2000).

Moreover, the fragmented study of individuals does not provide a viable basis for educational practice in current teaching and learning settings. As Shweder and Sullivan indicated, at least part of the current intellectual climate is "disposed to revalue processes and constraints that are local, variable, context-dependent, and in some sense made up" (1993, p. 502).

Thus, current influences on the field suggest that educational psychologists are becoming interdisciplinary in their perspectives in at least three broad domains: (a) in psychol-

Requests for reprints should be sent to Carol Kehr Tittle, Graduate School and University Center, City University of New York, 33 West 42nd Street, New York, NY 10036.

ogy—across the areas of psychology; (b) in the major subject matter disciplines—those who are involved in educational practice (e.g., the epistemology and pedagogy of mathematics, including psychology when constructs such as metacognition, self-regulation, and motivation are the goals of practice); and (c) in the domain of other disciplines involved in research in education—disciplines such as sociology, history, political science, anthropology, philosophy, and linguistics, among others that contribute to understanding and thinking about schooling, communities, and education. These ideas are consonant with comments by Wittrock (1992), Calfee (1992), and others on broadening methodologies and substantive areas in educational psychology.

Given these current directions and influences in the field, the remainder of this article identifies the major areas relevant to an educational psychology of assessment for practice. I propose a conceptual framework for an educational psychology of assessment for teaching and learning and briefly describe its major features. A particular context is used to examine the relevance of such a framework. The context is research examining teacher understanding of student responses to an assessment for use in junior high school mathematics classrooms. In the final section, I consider necessary developments in arguments for assessments that are valid in the context of practice, arguments grounded in an interdisciplinary educational psychology of assessment for teaching and learning.

ASSESSMENT: A FRAMEWORK FOR EDUCATIONAL PSYCHOLOGISTS

A framework to guide educational psychologists, and others, in thinking about assessments in research, development, and practice can be described along three dimensions, with a number of illustrative broad categories in each dimension. The three dimensions are epistemology and theories, the interpreter and user, and assessment characteristics. The illustrative, broad categories identified for each dimension are listed here (and described subsequently):

Dimension 1. Epistemology and theories
- Teaching and learning
- Curriculum
- Development and change

Dimension 2. Interpreter and user
- Knowledge
- Beliefs
- Intents and actions

Dimension 3. Assessment characteristics
- Embeddedness in practice
- Format and mode
- Scoring, evaluating, preparing, and feedback

The framework differs from other discussions of validation theory (e.g., Messick, 1989) by placing the interpreter and user as a central dimension. Interpreters and users are recognized as equal in importance to the dimensions of theory (typically identified as constructs) and assessment characteristics in the development, study, and evaluation of assessments.

The equal status given to the interpreter and user, to theory, and to assessment characteristics in an educational psychology of assessment is responsive to the various cognitive constructivist (and interpretive; Moss, 1992) perspectives on teaching and learning. The user or users will typically be teachers and students. Interpreters and users may also include parents or other school staff who construct interpretations and use assessments intended to further purposes of teaching and learning.

The framework can be used, in principle, whether assessments are developed by teachers or other external developers, and whether they are integral or separate from instructional activities. The framework is intended to suggest the lines of inquiry and data to be examined in the context of practice. These lines of inquiry and data form the basis of validation arguments (Cronbach, 1988; Shephard, 1993).

As an example of the lines of inquiry suggested by the framework, in the assessment example (given subsequently) think-aloud protocols are used to begin to understand and illustrate the intersections of the dimensions of the framework. In particular, for a student, there is an examination of interpretations of feeling (anxiety) in statements using different classroom activity statements as the context. For a teacher, there is an examination of the interpretations of student responses to such statements in the context of other knowledge and beliefs about a student and interpretations given for teaching actions. For the teacher example, views or theories of teacher development and change provide an example of the extension of validation arguments for assessments into a practice-based context.

Epistemology and Theories

The dimension of epistemology and theories can be illustrated with several categories. Theories of the grounds of knowledge, of teaching and learning, of curriculum, and of development and change can be identified for assessments and for interpreters/users. This is particularly true if an assessment is described as an assessment "system," allowing for examining the full context of assessment—preparation, process, outcomes, feedback, interpretations, and use. Thus, particular epistemologies and theories are explicitly or implicitly "held" by the interpreter/user and held or embodied by the processes and characteristics of assessments themselves. An educational psychology of assessment will include studies to understand these beliefs of interpreters and users and to evaluate the relationships and influences of the characteristics of assessments on these beliefs.

Teaching and learning. Educational psychologists are conversant with theories of teaching (Shulman, 1986), in-

struction, and learning. For example, Snow (1989) proposed a comprehensive list for assessing areas of student learning and performance (see Snow & Swanson, 1992). The area of individual development is also familiar, particularly as concerned with general patterns of development and mechanisms of change (e.g., Olson & Astington, 1993). Other, related topics such as motivation, individual differences, and social/personality theories can be readily identified. These theories or beliefs of teaching and learning are held by teachers, students, and, as Shephard (1991) pointed out, by developers of assessments.

Curriculum. This category is very broad and may be less familiar to educational psychologists. It includes the subject matter discipline and its embodiment in a school's and teacher's (and students') formal and informal or hidden curriculum (e.g., Cohen & Spillane, 1992; Gehrke, Knapp, & Sirotnik, 1992). For mathematics, the curriculum, theories of teaching and learning, and the development of mathematical thinking are at least partially exemplified by standards for curriculum and evaluation of the NCTM (1989), including the standards for teaching (NCTM, 1991). In specific instances, curricula are identified through local documents and practice.

Typically, curricular documents include both cognitive and affective/motivational considerations. For example, the NCTM standards (1989) include topics such as ways to develop mathematical reasoning, problem solving, communication, and connection of mathematical ideas. They also include mathematical dispositions—not simply attitudes, but a tendency to think and act in positive ways, which the standards suggest are manifested in ways students approach tasks and in a tendency to reflect on their own thinking (p. 233).

Concerns for equity and fairness for women and minorities are included in the curriculum area. Multicultural concerns in terms of representation and equity are also addressed through assessments as they reflect curricula.

Development and change. Theories of development and change include but are not limited to psychological theories of individual development and learning. Theories of how development and change occur for groups of students, for classrooms, or for groups of teachers (even for schools and districts) are assumed by the use of assessments. Examples are in the proposals for national tests and the use of tests by states and local school districts where the purpose is accountability or to direct instruction toward specific goals and methods. The effects of assessment for such purposes have begun to be studied more systematically in the past decade (Ellwein, Glass, & Smith, 1988; Haladyna, Nolan, & Hass, 1991; Shephard, 1991; Stake & Rugg, 1991; Tittle, Kelly-Benjamin, & Sacks, 1991). There is a longer tradition of examining the motivational and other effects of classroom tests and grading practices on individual students (e.g., Crooks, 1988; Natriello, 1987), although there are fewer data on student interpretations and uses of other assessments, including standardized test information.

There is also literature in program evaluation and organizational change that can be usefully related to assessments as part of educational practice. Sirotnik (1987), Cronbach et al. (1980), Goldenberg and Gallimore (1991), and Weiss (1992) provided ways to think about change. Sirotnik, for example, argued for local evaluation (and assessment) and change. These different perspectives can help in analyses of theories of change implicit in assessment characteristics and in those held by interpreters and users. They can usefully be made explicit and are likely to lead to greater tentativeness on the part of claims made for assessments; they may also lead to greater involvement of teachers in thinking about the processes of change and the design and use of external and classroom assessments.

Theories of change are likely to be different, depending on the interpreter (whether, e.g., a teacher or a student). Perspectives on teacher understanding of student development are emerging in the mathematics education research community (see, e.g., National Center for Research in Mathematical Science Education [NCRMSE], 1991, 1992; Schifter & Simon, 1992; Simon & Schifter, 1991 for descriptions of recent work). Most recently, work in this area has further expanded and includes attempts to describe change for groups of teachers within a school (e.g., Borko, Flory, & Cumbo, 1993).

Interpreter and User

The critical dimension for educational psychologists and others concerned with assessments in the teaching and learning context focuses on individuals who are interpreters and users. At minimum, these interpreters and users include teachers and students and may include other school or district professionals and parents. A cognitive constructivist perspective on assessment (Tittle & Forehand, 1992) suggests that teachers and learners construct schemas or integrate representations from assessments into existing views of the self, of teaching and learning, and of the curriculum, broadly construed. These interpretations include knowledge and beliefs and may also result in intents to use and actual use of assessments. A teacher may plan instructional activities or a student may increase time on a homework assignment.

The teacher as interpreter and user is examined in more detail here to illustrate the areas with which an educational psychology of assessment will be concerned. Categories or areas of representations in which teachers might find meaning or construct interpretations or narratives (Bruner, 1986) can be identified (see Tittle, 1989, 1991) and are shown in Table 1. Similar general categories can be noted for learners (e.g., Forehand, 1991; Snow, 1989) and could be developed and specified for other users of assessments among school staff and for parents (for parents, see Shephard & Bliem, 1993).

TABLE 1
Assessment Interpreter and User Dimension: The Teacher as Example

	Teacher Knowledge and Beliefs	*Teacher Intents and Actions*
Related to teaching	About the learner	Changes in practice
	Self-knowledge	Professional development
Related to assessment	Effects on students	Changes in planning
	Effects on teachers	Other assessment tasks

Teacher knowledge and beliefs. At least two broad areas of meaning can be identified in relation to assessment: (a) knowledge related to teaching and (b) knowledge about the assessment process. Knowledge related to teaching includes the teacher's knowledge about the learner and the teacher's self-knowledge. Teacher knowledge about the learner can include several domains. A teacher may construct "knowledge" about the learner's representations and conceptions about the subject matter (Confrey, 1990). Teachers may also identify learner beliefs about the self, the teacher, and learning. These learner beliefs may include such aspects as strategies, metacognition, self-regulation, and affective, motivational, and attributional beliefs. They may also include a learner's interpretations of teacher expectations, such as standards, beliefs about the subject matter and its epistemology and worth, and so on.

Teacher knowledge about the self as a teacher encompasses a spectrum of teacher cognitions and beliefs (Grossman, 1992; Kagan, 1990, 1992) about the self as teacher and about craft knowledge. For example, teachers may construct representations about the effectiveness of their own teaching with regard to content-specific pedagogical knowledge and strategies for different learners, about the effectiveness of their teaching with regard to motivational devices, and about their effectiveness as a teacher in general (Kagan, 1990; Shulman, 1986). A teacher may also make interpretations about local school and professional expectations, standards and values, and hence of her or his own effectiveness as a teacher.

Briefly, teachers may also have or construct further representations related to the assessment process itself. Teachers are likely to hold beliefs about effects of assessments on students before assessment (e.g., as providing a focus for learning), knowledge about assessment effects on students during assessment (e.g., as providing a sense of accomplishment or challenge or failure or inadequacy), and knowledge about assessment effects on students after assessment (e.g., as fair, meaningful, useful, as providing information for continuing development or lack thereof). Teachers may also hold beliefs about the effects of assessments on teachers themselves, such as requiring instruction on particular topics or problems or as providing (or not providing) useful information for instruction.

Teacher intents and actions. Assessments are also likely to have effects in the form of intents and actions. Teachers can form intentions to change classroom activities or projects or to interact differently with particular students. Teachers might take actions in the form of instructional planning or changing classroom interactions. They might critique and adapt classroom activities or design new instructional materials. Teachers might also take actions related to professional development, seeking additional or different interactions with colleagues, or undertaking further education activities for themselves.

Briefly, teachers might also take actions related to the assessment process. For example, before administering an assessment, teaching may be focused on particular subject matter topics or problem solving. After assessments, teacher actions might include additional probes or tasks to confirm or disconfirm interpretations formed on the basis of an assessment.

Assessment Characteristics

This dimension is concerned with the characteristics of assessments in the context of practice. In the development and evaluation of assessments, attention has often focused on format (structured/open) and mode of administration and response, as well as scoring. When considered in the context of teaching and learning practice, other characteristics become equally relevant. These characteristics include the degree to which assessments are "embedded" in ongoing instructional activities and descriptions of the complete system or program of assessment. The "system" characteristic refers to the activities surrounding the assessment—the preparation for assessment, the performance of the assessment task, its evaluation, and the description or reporting of assessment results, all of which contribute to the interpretation of assessments.

Embeddedness. Assessments can vary in the degree to which they are embedded in ongoing classroom instructional activities. At one extreme, assessment information is drawn from a typical classroom instructional activity, such as in laboratory experiments, writing assignments, or other settings in which the processes and outcomes are perceived by both teacher and student as a part of ongoing learning activities. Teacher tests, special activities, or assignments designated for evaluation are a part of many classrooms, yet the assessment is not as directly a part of ongoing learning

activities. At the other extreme are assessments that are externally designated for classroom administration or those that may be administered by a special proctor, not the classroom teacher.

Mode and format. Assessments differ in mode and format. Assessment "stimuli" may be embodied in a variety of modes: print, auditory, audiovisual (perhaps videotape), real-time laboratory, records of observations, and group and class interactions, among others. Assessments can also elicit or specify a variety of response modes: oral, written documents, videotape, artistic performances/products, and so on. Assessment and response formats can also vary, from detailed and highly structured to broadly described and open in structure. Examples are selecting or constructing a response, from multiple-choice responses to preparing drawing or writing portfolios.

Scoring/evaluation. Scoring is a function of the response mode, to some degree: from checking against a predetermined key to holistic, impressionistic ratings. In between are various versions of detailed marking schemes or rubrics that combine holistic and focused scoring (e.g., the Quasar mathematics assessment; Silver & Cai, 1993). Scoring or evaluation may or may not involve comparisons with other students or standards of performance. Scoring and evaluation may or may not involve the teacher who provided or administered the assessment task. It may or may not include attempts to moderate (develop or monitor, or both, common standards in marking) assessments across schools or other educational units for external assessments or among teachers in a school for local assessments.

Preparation/description/reporting. This category calls direct attention to the immediate context for an assessment—the system of which the assessment performance is the (nominal) central event. Preparations for assessments can vary from almost none—a teacher may remind students to review particular topics—to very extensive test-preparation courses and coaching. External testing programs may involve elaborate preparation for teachers and students by using practice assessments, test-preparation materials, and workshops for teachers. Reporting, feedback, and descriptions of performance can vary from a teacher's checkmark or written comments on an essay, or other assessment, to elaborate systems to provide interpretive comments to teachers or students, or both, using computer-based algorithms and expert interpretations. Such reporting or interpretations may or may not include connections to curricula or to instructional suggestions.

In summary, I have identified three dimensions (theories and epistemology, interpreters and users, and assessment characteristics) to provide the basis for an educational psychology of assessment. The assessment characteristics are critical for their relationship to theories and change and to the interpretations made about them by students and teachers, in the context of teaching and learning. To illustrate implications of the framework, I describe a specific assessment intended for classroom use. As the assessment is described, aspects of the framework can be identified. In the description I also illustrate one approach to thinking about teachers as interpreters and users of assessments.

AN EXAMPLE: RESEARCH ON TEACHER UNDERSTANDING OF A PARTICULAR ASSESSMENT

The Mathematics Assessment Questionnaire

The assessment used as an example here is the Mathematics Assessment Questionnaire (MAQ). The purpose of the MAQ is to provide information to teachers on student thoughts and feelings about solving mathematical word problems, for students in Grades 7 through 9. The assumptions or views of the learner and teacher are that students are active constructors of meaning and can directly reflect on their thinking and beliefs; that student learning of mathematical knowledge occurs in the context of student thoughts and feelings about mathematical knowledge; and that teacher understanding of student beliefs can facilitate student development of mathematical knowledge.

The questionnaire surveys students' thoughts and feelings about learning and doing mathematical word problems in the context of three classroom activity settings (the full rationale and development research are described in Hecht & Tittle, 1990; Tittle & Hecht, 1990).[1] In general, the questionnaire

[1]Psychometric characteristics of the Mathematics Assessment Questionnaire (MAQ) are briefly described here. The questionnaire was developed through a series of pilot and try-out studies. There was a large-scale administration of the MAQ in a paper-and-pencil version with about 1,500 students in Grades 7, 8, and 9 in eight urban public schools. Analyses of variance (ANOVAs) by grade and sex indicated few item differences, and the statistical studies of the MAQ were carried out for the total sample. Data for the metacognitive and self-regulation statements are reported only for individual statements. However, factor analyses of these statements are available and suggest a two-factor solution for the metacognitive statements, awareness of monitoring and checking of the problem-solving process and awareness of general problem-solving approaches and strategies. A multidimensional scaling analysis of the self-regulatory statements across the three activity statements provides some support for the activity structure used in writing items (Tittle, Weinberg, & Hecht, 1994) and indicates that students perceive the settings as somewhat different psychologically. Factor analyses of the self-regulatory statements are available also in Hecht and Tittle (1990). Two- or three-factor solutions were obtained for the three activity settings. Coefficient alphas were used to obtain internal consistency estimates for the three-item clusters of the 10 affective, motivational, and attributional constructs in each activity setting. The reliability estimates were consistently highest for the Homework setting (.48 to .74, median = .61); the least consistent estimates were for the construct External Performance Goals, .20 in During Class and .53 and .54 in Working With Others and Homework.

statements ask students to reflect on learning and doing mathematical word problems to assess two general areas of student characteristics: (a) awareness of self-regulatory skills and beliefs; and (b) affective, motivational, and attributional beliefs.[2] Statements include a context, one of three classroom activities: (a) during class, when a teacher leads a lesson (or alternatively, a student-centered classroom discourse model, based on the NCTM, 1991); (b) working with others, in a problem-solving group; or (c) doing homework, an independent activity. Sample statements illustrate the facet approach used to develop statements:

> *Self-regulation:* During Class—When I can think of another way to solve a word problem, I volunteer to show the class;
>
> *Anxiety:* Working with Others—I dread the thought of trying to solve a math word problem with other students;
>
> *Internal learning goals:* Homework—I like to do hard homework math word problems because I learn more math by working them.

Figure 1 provides the general outline of the 143-item questionnaire.

Student interpretations of statements. As one example of understanding the meanings that students may construct from an assessment, several teachers administered one set of metacognitive statements, which asked students what they do before, during, and after solving a nonroutine mathematics problem. Students first solved a nonroutine word problem and then responded to statements asking them to reflect on what they did when they solved this particular problem. Students responded yes, no, or maybe I did this, to statements such as, "I looked back at the problem to see if my answer made sense" and "I drew a picture to help me understand the problem."

In several classes we asked students to turn over the page of statements and answer this question after they had worked the nonroutine mathematics problem: "How did thinking about these questions help you think about how you do math problems?"

Sample student responses include: "It gave me more ways to work a problem," "Answering these questions made me think about how I solved the problem and made me double-check my answer," "It helps me think about what to do whenever I have math problems, and it also helps me recall whether I check my work or not," "Answering these questions helps me think about things that I don't do on math problems," and "The questions didn't really help me think about how you do math problems. But it gets you thinking."

Another procedure used was to ask students to talk aloud as they considered questionnaire statements. This procedure also provides examples of beginning to understand the meanings that learners construct from assessments. Here is an example of a student talking about anxiety or concern statements in each of the three different activity settings (using a computer-administered version of the questionnaire):

DURING CLASS
Statement:
I am afraid when I have to ask my math teacher a question about a word problem during class.
Student response:
"afraid. Well, sometimes I'm like afraid because if everyone else seems like they have gotten it and like I'm totally lost it could be kind of embarrassing. That used to happen, but not always, so it is sort of true ... sometimes if I'm just like lost I don't want to raise my hand or anything because I am afraid I'll be branded. If it is important and I don't understand it, like it is the day before a test, then I will just raise my hand ... so I guess it is sort of true."
WORKING WITH OTHER STUDENTS
Statement:
I dread the thought of trying to solve a math word problem with other students.
Student response:
"not at all true, because you can learn more that way."
HOMEWORK
Statement:
I feel nervous when I think about doing hard word problems for homework.
Student response:
"... I feel kind of nervous ... cause I want to show that at least I can do good effort."

Teacher Program Disk

The questionnaire was originally developed in paper-and-pencil format. Programs have been developed so that the student survey can be computer-administered and the results

Overall, the median reliability of the three item clusters was .60. The desire to sample a broad spectrum of constructs resulted in using these small sets of items. The computer-based *Teacher Program* (Tittle & Hecht, 1994) consistently presents teachers with statements and student responses, however, focusing attention at an appropriate level for interpretation.

[2]Definitions of the current constructs in the MAQ are as follows: Metacognition: self-awareness of activities/strategies when working a nonroutine word problem; Self-regulation: awareness of self-directed strategies used to learn, in classroom learning contexts, for math word problems; Value: the conviction that learning about math word problems is worthwhile or useful; Interest: liking to work or learn about mathematical word problems; Confidence: belief in one's ability to do or learn mathematical word problems; Anxiety: worry, uneasiness, or fear about doing or learning mathematical word problems; Internal learning goals: internal or intrinsic motivations for learning math word problems; External performance goals: external motivations for learning math word problems (e.g., grades, approval); Internal stable controllable: attribution of success (During Class, Homework) or failure (Working With Others) to an internal and changeable cause (effort); Unknown control: perceived confusion or inability to make sense out of why one succeeds or fails at doing mathematical word problems.

CONSTRUCT	ACTIVITY SETTING		
Metacognitive: Solving a problem .Planning, define objectives .Monitoring progress .Checking & evaluating .Strategies employed	METACOGNITIVE ITEMS NOT LINKED TO SPECIFIC SETTING LINKED TO SPECIFIC PROBLEM		
CONSTRUCT	During Class	Working With Others	Doing Homework
Self-Regulation			
Affective Beliefs			
.Value			
.Interest			
.Confidence			
.Anxiety			
Motivations			
.Internal Learning Goals			
.External Performance Goals			
Attributions			
.Internal Stable Controllable			
.Unknown Control			

FIGURE 1 Domain specifications for the Mathematics Assessment Questionnaire (MAQ)—two facets: psychological construct and setting.

compiled on a teacher program disk (Tittle & Hecht, 1994). The computer version supports teacher examination of student responses, including a notebook feature that encourages students to write to their teacher, and a help feature that includes descriptions of available reports and suggested instructional strategies for all sections of the questionnaire. The teacher program also provides a research tool for beginning to describe how teachers interpret and consider use of this particular assessment. Before describing how teachers begin to think about student responses, it is important to illustrate the screens the teacher sees when reviewing student questionnaire responses.

Teachers can access student responses at a class level and a student level. At the class level, the names of students selecting each alternative are immediately available. Examples of class-level responses to self-regulation statements (During Class) and value statements (During Class) are given in the two screen prints of Figure 2a and 2b.

Although the self-regulation and metacognitive statements are interpreted only at the individual statement level, the remaining affective, motivational, and attributional beliefs comprise three-item clusters. These clusters of statements can be reviewed and interpreted at the individual statement level or by "Possible Need" indicators. These latter indicators are similar to criterion-referenced indicators: The criterion that is used requires that the student has selected two of three or three of three responses at the two extreme positions of the response scale in a direction that suggests a possible need for follow-up (e.g., lack of confidence, high anxiety). A student's responses are said to indicate a "Possible Need for Follow-Up" when they meet the criterion. This indicator is very directly interpreted, as the definition suggests.[3]

The need indicators are available for a class, summarized, and also for individual students. Figure 3a illustrates the student-level screen for the Needs for Jennifer, and Figure 3b for Matthew, both seventh-grade students. Note that the

[3]The development and data on the need indicators are described in Hecht and Tittle (1990).

SUMMARY OF CLASS RESPONSES TO SELF-REGULATION STATEMENTS
(N=21)

DURING CLASS

During a lesson:	VERY TRUE	TRUE	SORT OF TRUE	NOT VERY TRUE	NOT AT ALL TRUE	NO RESPONSE
10. I think about whether I understand an example the teacher puts on the board.	6 29%	8 38%	5 24%	1 5%	1 5%	0 0%
11. When my math teacher makes a mistake, I say something about the error.	8 38%	4 19%	3 14%	3 14%	3 14%	0 0%
12. I ask my math teacher to explain a problem again that I do not understand.	6 29%	5 24%	5 24%	4 19%	0 0%	1 5%

(Percentages based on students answering problems, including no responses.)
F1 Help | P Print | ↑↓ Bar | ↵ Students | PgUp/PgDn Questions | TAB Settings

FIGURE 2(a) Sample screen print of class responses to self-regulation statements: During Class.

SUMMARY OF CLASS BELIEFS ABOUT WORKING WORD PROBLEMS
(N=21)

DURING CLASS

VALUE	VERY TRUE	TRUE	SORT OF TRUE	NOT VERY TRUE	NOT AT ALL TRUE	NO RESPONSE
24. Even when I listen to my teacher, I cannot understand how learning to solve word problems will help me in my everyday life.	7 33%	0 0%	7 33%	3 14%	4 19%	0 0%
26. It is important to learn to do the types of word problems my teacher explains in class.	9 43%	6 29%	4 19%	0 0%	2 10%	0 0%
31. Listening to my math teacher explain word problems during class helps me see how important math is.	5 24%	6 29%	8 38%	0 0%	2 10%	0 0%

(Percentages based on students answering problems, including no responses.)
F1 Help | P Print | ↑↓ Bar | ↵ Students | PgUp/PgDn Beliefs | TAB Settings

FIGURE 2(b) Sample screen print of class responses to value statements: During Class.

screen includes the statements to which the student responded and the responses. Figure 4a illustrates the notes written by Jennifer when she was responding to one of the Anxiety statements (the program provides the teacher information about what was on the screen when the student opened her notebook). Figure 4b provides a note written by Matthew.

As these screen prints for individual students indicate, a teacher will have different sets of student responses within the same mathematics class, such as students who are anxious in the classroom setting (Jennifer) and those who indicate that working with others in problem-solving groups is of little value to them (Matthew). An educational psychology of assessment is concerned to go beyond analyses of student responses to understand how teachers interpret and use as-

sessments. The framework proposed for assessment requires working with teachers to develop such an understanding.

Teacher Levels of Understanding of Student Responses

The teacher program offers opportunities to examine the manner in which teachers explore unfamiliar assessment information and information made available in an unfamiliar computer mode. From the perspective of an educational psychology of assessment, it is important to explore the meaning and uses teachers construct from assessments. Also, in this broader context, it is important to consider teacher change and to begin to describe teacher change and teacher development (cf. Fuller, 1969) as a function of experience with assessments.

The method used with teachers is based on think-aloud

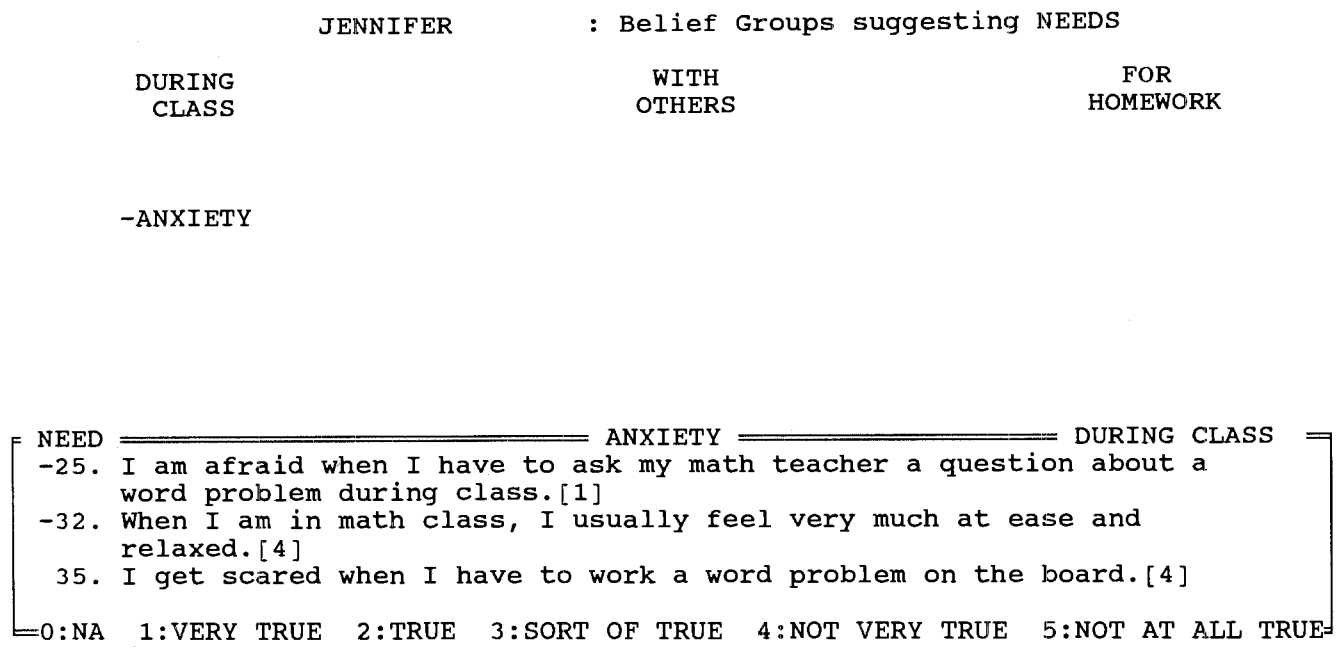

FIGURE 3(a) Sample screen print of student NEEDS: Jennifer.

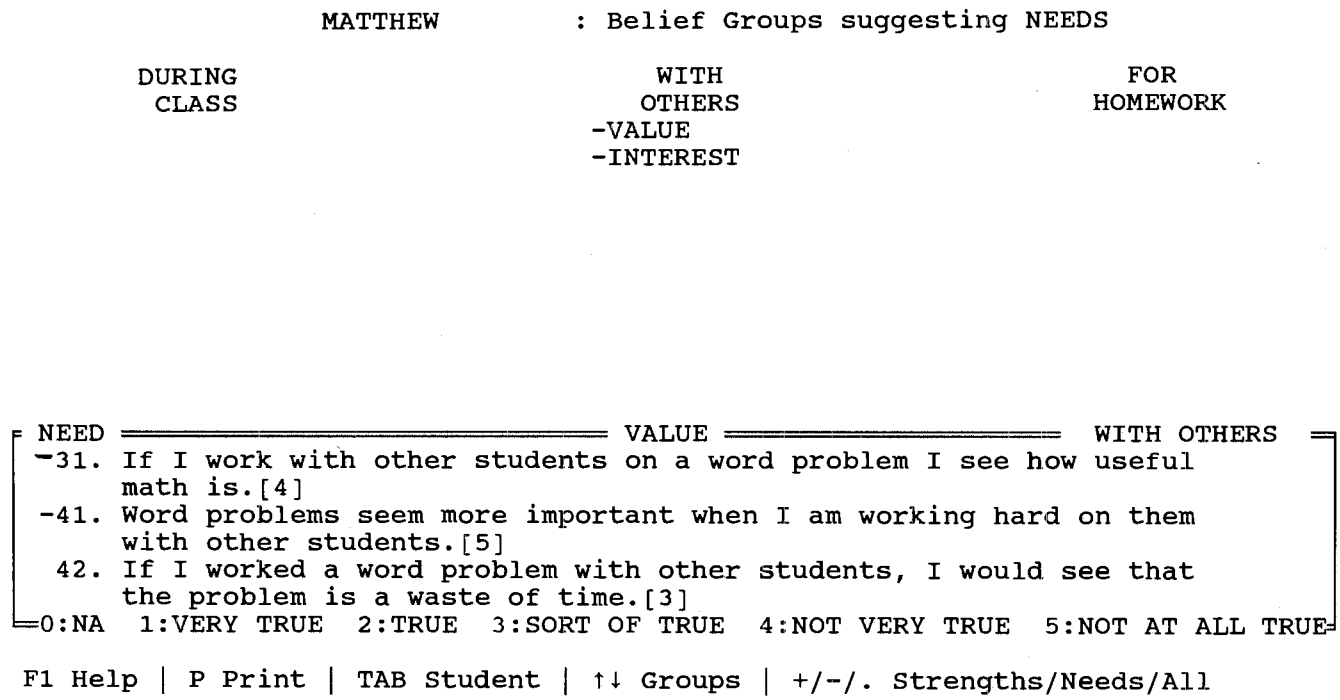

FIGURE 3(b) Sample screen print of student NEEDS: Matthew.

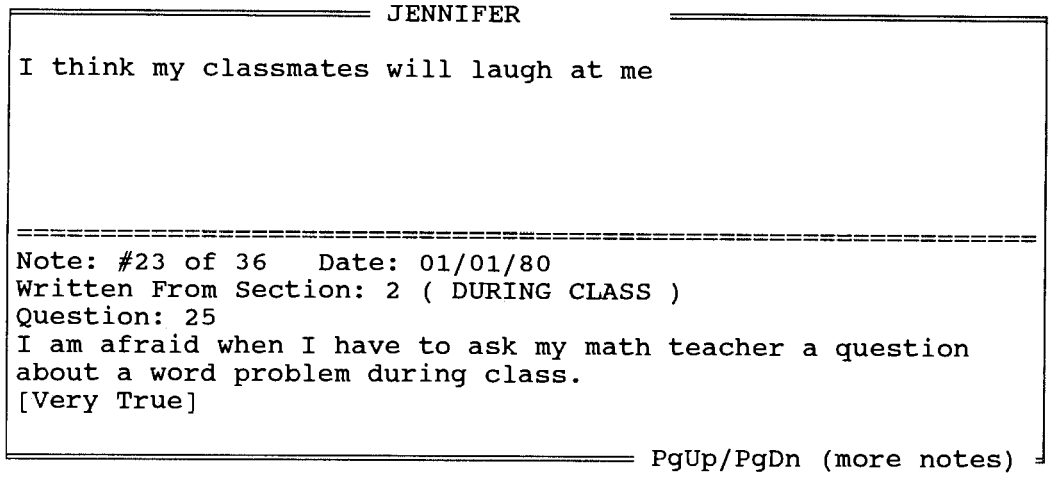

FIGURE 4(a) Sample NOTE print: Jennifer, Statement 25.

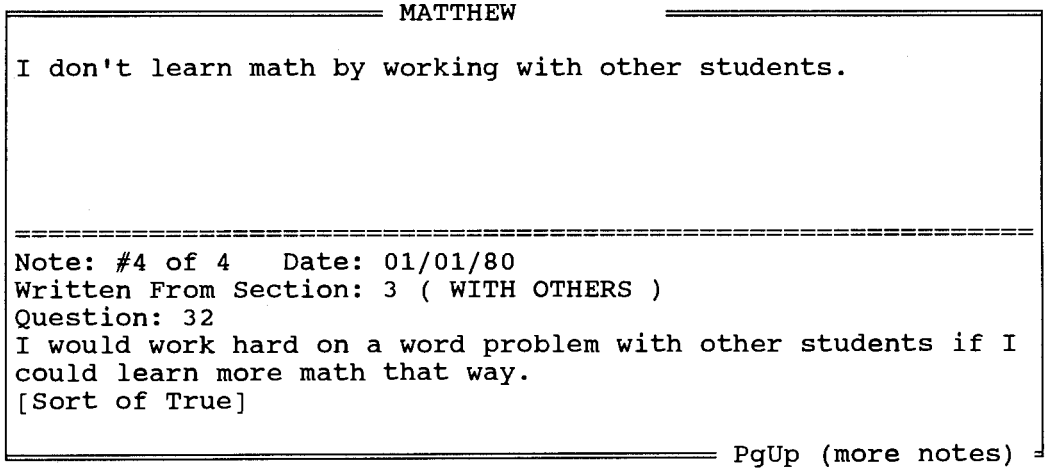

FIGURE 4(b) Sample NOTE print: Matthew, Statement 32.

procedures (Ericsson & Simon, 1984). In this procedure, the teacher talks aloud as he or she explores the structure of the program, the structure of the assessment, the psychological constructs in the survey, and the responses of his or her class and individual students.

We have developed a provisional framework to describe teacher change using this particular assessment—the MAQ. Four levels of understanding are considered, levels of the teacher's constructing meaning from the assessment information, using the teacher computer-based program with class-level and individual-level student responses. The four Levels of Teacher Understanding proposed for the computer-based version of the MAQ are

Level 1 Acquiring procedural skills and conceptual structures:

a. Facility with the computer program and general structure of assessment information.

b. Understanding constructs (definitions) of assessment.

Level 2 Contextualizing student responses in the psychological domain:

a. Accessing other, relevant information about the student.

b. Interpreting student or class response conditional on this other information.

Level 3 Using the contextualized information to

select or develop specific instructional strategies.

Level 4 Internalizing and transforming the assessment information about the psychological domain into other instructional settings and practices.

Using typed audiotape transcriptions from the talk-aloud sessions, teacher statements for the first three levels have been identified. (Level 4 remains speculative.) An example of a Level 1 statement acquiring procedural skills and conceptual structures is, "I like the summary of their responses. Where can I get that? I like to see the needs and strengths ... I'm looking at the summary chart." An example of a Level 2 statement contextualizing the students' responses in the particular psychological domain is, "Julie ... I wouldn't suspect that she didn't like working with other students, but now that I think of it, she is one who likes to stay at her own desk in another part of the room." An example of a Level 3 statement using information from the MAQ to select or develop specific instructional strategies is, "Confidence ... she probably doesn't like it when I put her in a group which is why she's withdrawn and that would either make me want to pair her up with somebody that could help her with her anxiety about this or make sure that I give her individual attention." In Levels 2 and 3, the teacher's statements are related to other information she has about the student, reflecting on the information and going on to propose specific instructional strategies.

The Levels of Understanding proposed are based on work with a small number of teachers (five), and we are working on additional teacher transcripts. The framework illustrated here does not yet reflect a more "fine-grained" analysis of the protocols from the teacher talk-alouds. The next steps in analysis of the transcripts are for categories of teacher understanding and use. Several questions will be examined, including, "What is the pattern of levels within one teacher transcript? Across transcripts of different teachers? Does the pattern of levels change if a teacher uses the questionnaire more than once?" Other questions suggested by the particular assessment used here are, "What psychological constructs do teachers select spontaneously when reviewing student responses? What emphasis is given to class-level versus student-level responses?"

The research on teachers and assessment also needs to be related to work on teacher change in the mathematics-education community. The levels proposed here began from the work of Hall and Loucks (1977) in evaluation. In the mathematics education research community, Schifter and Simon (1992) and Franke, Fennema, Carpenter, and Ansell (1992) have proposed frameworks for teacher development. Schifter and Simon are concerned with understanding the reasons teachers give for particular instructional activities—that they focus on student learning from a constructivist perspective and, more recently, on teachers' understanding of mathematics (D. Schifter, personal communication, July, 14, 1993). Work at the Cognitively Guided Instruction project (Franke et al., 1992) focuses on three areas of teacher change: teacher knowledge of the content of children's thinking, teacher beliefs (that children can solve problems without direct instruction from the teacher), and teacher use of children's thinking in the classroom.

Table 1 describes the domains in which teachers as interpreters and users might represent assessment information. This broader, educational psychology framework for assessment is closer to the other teacher-change frameworks in mathematics education. Future research interviews may usefully be more detailed, adding structure to the open-ended format used to date. For example, other areas of teacher change can be explored, areas that more explicitly incorporate the underlying epistemology of the survey as both Schifter and Fennema have done in mathematics teaching. The expansion of the levels of understanding might look like Table 2. The adaptation draws on Franke et al. (1992) and Schifter and Simon (1992).

The development of Teacher Levels of Understanding of the MAQ illustrates the analyses and relationships an educational psychology of assessment for practice is likely to develop. The adaptation shown in Table 2 indicates a more direct focus on the teacher as assessment user and on classroom practice itself. Such analyses will encourage more direct links between instruction and assessment.

An additional dimension, not considered by the teacher-change frameworks described so far, is evident in the teacher transcripts. This dimension or strand is that of teacher-change processes, the self-knowledge and risk taking required for change. Teacher transcripts include statements that reflect a challenge to the teacher's beliefs about herself or himself as a teacher and the difficulty of changing, including taking risks. In the transcripts there are referrals to such issues: "Oh no (#44. I get bored when other students are working word problems on the board in math class) ..." and "... very true, what a stab wound to my heart." (CL:3/19/92, p. 3)

Another teacher raised a similar issue:

> I don't know why I'm so hung up on that, but there is a connection between interest and value ... Possible activities ... again these sound really good, it's just there's this constant pressure for the curriculum to get done but ... I would love to do some of these instructional strategies and maybe I just need to be brave and break away from the curriculum. (EF:11/24/92, p. 10)

Thus, in addition to a focus on changes in teacher beliefs and activities that are specific to the context of practice and assessment, educational psychologists may contribute to the interdisciplinary study of teachers and students through understanding of change processes themselves (e.g., Mahoney, 1991). The framework for assessment incorporates the change process itself, as indicated in the earlier discussion of the theories dimension of the framework. This category recognizes the meaning making of teachers and students with respect to their own change—as a proactive process that is self-organized (see also Schifter & Fosnot, 1993).

TABLE 2
Teacher Levels of Understanding and Beliefs About the Mathematics Assessment Questionnaire—Practice-Related[a]

Level	*Understandings and Beliefs*
1	Acquires procedural skills and conceptual structures: Computer program, assessment information, psychological constructs (mathematical dispositions)
2	Contextualizes student responses in the psychological domain: accesses other relevant information; interprets responses conditional on this information
3	Uses contextualized information to select or develop classroom instructional activities
4	Hold beliefs about students' active construction of beliefs (mathematical dispositions) and modifiability of these beliefs in the context of learning mathematics: practice focused on teacher
5	Holds beliefs; uses student beliefs (affective, etc.) in classroom: practice focused on students • Focuses on providing opportunities for students to solve problems and talk about their monitoring and affective (and other) beliefs and feelings and • listens to students talking with this framework and uses what is heard to make instructional decisions

[a]Adaptation also draws on Franke, Fennema, Carpenter, & Ansell (1992) and Schifter & Simon (1992).

CONCLUDING COMMENTS: ON VALIDITY AND THE FRAMEWORK

Before drawing implications for validation strategies, arguments, and evidence, there needs to be an acknowledgment of the situational specificity of the work described. The general descriptive framework for Levels of Teacher Understanding and Beliefs are particular to the MAQ. However, this limitation does not preclude drawing implications for test theory and validity theory.

Validity and test theory are not directly identified in the broadly based, three-dimensional assessment framework described earlier: epistemology and theories, interpreter and user, and assessment characteristics. Validity theory deals at a general level with inferences and appropriate uses of assessments. However, validity theory as currently formulated does not lead directly to arguments (Kane, 1992; Shephard, 1993) for studies of the meaning and use of assessments by the several groups of users of assessments in education. In this assessment framework, such studies are needed throughout the assessment-development process as well as when assessments are implemented and used in practice. Assessments must have meaning for students and teachers in relation to teaching and learning. Thus, from the standpoint of an educational psychology of assessment, current discussions of test theory (Mislevy, 1993) and validity (Messick, 1989) are inadequate because they do not take a cognitive and constructivist perspective on the interpreter and user. These discussions focus on the scientific constructs with which the assessments are concerned and do not provide evidence on the inferences and uses made in the context of practice by teachers and learners themselves (although Linn, Baker, & Dunbar, 1991, acknowledge the role of the user in the performance assessment context).

Many of the critics of testing have pointed to the unintended consequences or outcomes of testing, such as coaching and teaching to the test. However, the terms *consequences* or *unintended consequences* do not reflect positively on the world of practice. There is a need for more positive, proactive models of the relationships between practice and assessment theory.

An analogy to medical clinical trials may be helpful. Drugs are interventions on the human system and are not introduced without clinical, or limited, trials in practice. Although such trials may not uncover all misuses or side effects of the drug, they provide much useful information. In some ways such clinical trials can be described as positive, proactive collaborations of researchers/developers, practitioners, and patients.

In educational settings and systems, assessments are typically introduced without clinical trials. It is argued here that studies of the effects of assessments on teachers, students, and parents, among others, in limited, practice-based settings are important for evaluating (providing validity-related evidence for) assessments. In the best situations, such studies would be collaborations with practitioners (teachers) and students. In situations in which assessments are developed by practitioners, it might be beneficial for "external" collaborators to participate—discipline-based specialists, educational psychologists, or others. The work reported by Shephard and her colleagues (e.g., Flexer & Gerstner, 1993) may illustrate one prototype.

The expanding of educational psychology, both from interdisciplinary and practice-based orientations, suggests that teachers (practitioners), subject matter specialists, educational psychologists, and others concerned with assessment for teaching and learning can collaborate. The collaboration and colleagueship goes both ways—from external initiatives to include work with those in classroom practice and from classroom practice initiatives to include work with those who are external.

The framework proposed here for an educational psychology of assessment begins to specify what needs to be examined in practice-based trials of assessments. Theories and

epistemologies of science and practice underlie and are basic to validation arguments. Both test theory and validity theory need expansion to encompass the world of practice-based assessments, to encompass the several epistemologies and theories of teaching, learning, curriculum, and change, and to view the interpreter and user as central to the development and evaluation of assessments.

Educational psychology has new challenges to define proactively what it means to develop and evaluate assessments for practice—that is, assessments that support teachers and students in constructing meanings that have use for them in their educational situations. Major steps have been taken toward putting psychology into educational measurement (e.g., Messick, 1984, 1989; Snow, 1989). The next step is to integrate educational phenomena and practice into the development, trials, and evaluation (validation) of educational assessments. Meeting these challenges requires developing new principles for a true educational psychology of assessment for teaching and learning.

ACKNOWLEDGMENTS

This article is a revised version of the Presidential address for Division 15, Educational Psychology, of the American Psychological Association, presented at the annual meeting, Toronto, Canada, August 1993.

This work was partially supported by the Ford and Aaron Diamond Foundations.

REFERENCES

Berliner, D. C. (1992, August). *The science of psychology and the practice of schooling: The one hundred year journey of educational psychology from interest, to disdain, to respect for practice.* Paper presented at the meeting of the American Psychological Association, Washington, DC.

Borko, H., Flory, M., & Cumbo, K. (1993, April). *Teachers' ideas and practices about assessment and instruction.* Paper presented at the meeting of the American Educational Research Association, Atlanta.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Calfee, R. (1992). Refining educational psychology: The case of the missing links. *Educational Psychologist, 27,* 163–175.

Carroll, J. B. (1993). Educational psychology in the 21st century. *Educational Psychologist, 28,* 89–95.

Cohen, D. K., & Spillane, J. P. (1992). Policy and practice: The relations between governance and instruction. *Review of Research in Education, 18,* 3–49.

Confrey, J. (1990). A review of the research on student conceptions in mathematics, science and programming. *Review of Research in Education, 16,* 3–56.

Cronbach, L. J. (1988). Five perspectives on the validity argument. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 3–17). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cronbach, L. J., Robinson, S., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation.* San Francisco: Jossey-Bass.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58,* 438–481.

Ellwein, M. C., Glass, G. V., & Smith, M. L. (1988). Standards of competence: Propositions in the nature of testing reforms. *Educational Researcher, 17,* 4–9.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis.* Cambridge, MA: MIT Press.

Flexer, R. J., & Gerstner, E. A. (1993, April). *Dilemmas and issues for teachers developing performance assessments in mathematics.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Forehand, G. A. (1991, April). *Assessment and the construction of meaning for instruction and learning: The learner's perspective.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Franke, M. L., Fennema, E., Carpenter, T. P., & Ansell, E. (1992, April). *The process of teacher change in cognitively guided instruction.* Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal, 6,* 207–226.

Gehrke, N. J., Knapp, M. S., & Sirotnik, K. A. (1992). In search of the school curriculum. *Review of Research in Education, 18,* 51–110.

Goldenberg, C., & Gallimore, R. (1991). Local knowledge, research knowledge and educational change: A case study of early Spanish reading improvement. *Educational Researcher, 20*(8), 2–14.

Grossman, P. L. (1992). Why models matter: An alternate view on professional growth in teaching. *Review of Educational Research, 62,* 171–179.

Haladyna, T. M., Nolan, S. B., & Hass, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20*(5), 2–7.

Hall, G. E., & Loucks, S. F. (1977). A developmental model for determining whether the treatment is actually implemented. *American Educational Research Journal, 14,* 263–276.

Hecht, D., & Tittle, C. K. (1990, June). *Mathematics Assessment Questionnaire: Technical report* (research ed.). New York: Center for Advanced Study in Education, Graduate School, City University of New York.

Kagan, D. M. (1990). Ways of evaluating teacher cognition: Inferences concerning the goldilocks principle. *Review of Educational Research, 60,* 419–469.

Kagan, D. M. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research, 62,* 129–169.

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112,* 527–535.

Klausmeier, H. J. (1988). The future of educational psychology and the content of the graduate program in educational psychology. *Educational Psychologist, 23,* 203–219.

Lesgold, A., & Glaser, R. (1989). *Foundations for a psychology of education.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15–21.

Mahoney, M. J. (1991). *Human change processes: The scientific foundations of psychological therapy.* New York: Basic Books.

Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement, 21,* 215–237.

Messick, S. J. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–303). Washington, DC: American Council on Education and the National Council on Measurement in Education.

Mislevy, R. J. (1993, June). *Test theory reconsidered.* Revised version of a paper presented at the annual meeting of the National Council on Measurement in Education, Atlanta, GA.

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research, 62,* 229–258.

National Center for Research in Mathematical Science Education. (1991). *Methodologies for studying teacher change in the reform of school mathematics.* Blacksburg, VA: Author.

National Center for Research in Mathematical Science Education. (1992). *Methodologies for studying teacher change in the reform of school mathematics: Vol. II.* Nashville, TN: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics.* Reston, VA: Author.

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist, 22,* 155–175.

Olson, D. R., & Astington, J. W. (1993). Thinking about thinking: Learning how to take statements and hold beliefs. *Educational Psychologist, 28,* 7–23.

Scheurman, G., Heeringa, K., Rocklin, T., & Lohman, D. R. (1993). Educational psychology: A view from within the discipline. *Educational Psychologist, 28,* 97–115.

Schifter, D., & Fosnot, C. T. (1993). *Reconstructing mathematics education: Stories of teachers meeting the challenge of reform.* New York: Teachers College Press.

Schifter, D., & Simon, M. A. (1992). Assessing teachers' development of a constructivist view of mathematics learning. *Teaching & Teacher Education, 8,* 187–197.

Shephard, L. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 20*(6), 2–16.

Shephard, L. (1993). Evaluating test validity. *Review of Research in Education, 19,* 405–450.

Shepard, L. A., & Bliem, C. L. (1993, April). *Parent opinions about standardized tests, teacher's information, and performance assessments.* Paper presented at the meeting of the American Educational Research Association, Atlanta, GA.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 3–36). New York: Macmillan.

Shweder, R. A., & Sullivan, M. A. (1993). Cultural psychology: Who needs it? *Annual Review of Psychology, 44,* 497–523.

Silver, E. A., & Cai, J. (1993, April). *Schemes for analyzing student responses to QUASAR's performance assessments: Blending cognitive and psychometric considerations.* Paper presented at the meeting of the American Educational Research Association, Atlanta, GA.

Simon, M. A., & Schifter, D. (1991). Towards a constructivist perspective: An intervention study of mathematics teacher development. *Educational Studies in Mathematics, 22,* 309–331.

Sirotnik, K. A. (1987). Evaluation in the ecology of schooling: The process of school renewal. In J. Goodlad (Ed.), *The ecology of schooling. NSSE Yearbook* (pp. 41–61). Chicago, IL: University of Chicago Press.

Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher, 18*(9), 8–14.

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology, 43,* 583–626.

Stake, R. E., & Rugg, D. (1991). Impact on the classroom. In R. G. O'Sullivan (Ed.), *Effects of mandated assessments on teaching* (pp. xix–xxii). Greenwich, CT: JAI.

Tittle, C. K. (1989). Validity: Whose construction is it in the teaching and learning context? *Educational Measurement: Issues and Practice, 8,* 5–13, 34.

Tittle, C. K. (1991, April). *Assessment and the construction of meaning for instruction and learning: The meaning and use of assessments by teachers.* Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Tittle, C. K., & Forehand, G. A. (1992). *Assessment and the construction of meaning for instruction and learning.* Unpublished manuscript.

Tittle, C. K., & Hecht, D. (1990, June). *Mathematics Assessment Questionnaire: A survey of thoughts and feelings for students in grades 7–9. Manual for users.* New York: Center for Advanced Study in Education, Graduate School, City University of New York.

Tittle, C. K. & Hecht, D. (1994). *Mathematics Assessment Questionnaire, Teacher Program Disk and Student Program Disk* [Computer program]. New York: Center for Advanced Study in Education, Graduate Center, City University of New York.

Tittle, C. K., Kelly-Benjamin, K., & Sacks, J. (1991). The construction of validity and effects of large scale assessment in the schools. In R. E. Stake (Ed.), *Effects of changes in assessment policy: Vol. 1. Advances in program evaluation* (pp. 233–254). Greenwich, CT: JAI.

Tittle, C. K., Weinberg, S. L., & Hecht, D. (1994). *A facet approach to affective measurement for classroom activity settings in mathematics.* Unpublished manuscript.

Weiss, C. H. (1992). Research should supplement, not supplant, lay inquiry [Review of *Inquiry and change: The troubled attempt to understand and shape society*]. *Educational Researcher, 21*(7), 38–40.

Wittrock, M. C. (Ed.). (1986). *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

Wittrock, M. C. (1992). An empowering conception of educational psychology. *Educational Psychologist, 27,* 129–141.