

THE PARADOX OF CLASSROOM ASSESSMENT: A CHALLENGE FOR THE 21ST CENTURY

STEVEN KATZ, LORNA EARL, & DAVID OLSON

Ontario Institute for Studies in Education/University of Toronto

ABSTRACT. There are at least two very different conceptions of competence at play on the landscape of formal education, best defined in terms of what is known collectively and what is known subjectively. Teachers are forced to attend to each of these in their challenge to meet both the demands of a fixed school curriculum and the needs of the individual children with their varied understandings, backgrounds, and interests who make up a class. We outline these two contrasting notions of competence here, including their associated epistemological, pedagogical, psychological, and assessment implications. This approach shows how classroom assessment is caught in a paradox: it attempts to serve competing purposes that are intimately bound to alternative conceptions of competence.

RÉSUMÉ. Il existe au moins deux conceptions radicalement différentes de la compétence dans le paysage de l'éducation formelle, que l'on pourrait définir sous l'angle de ce qui est connu collectivement et de ce qui est connu subjectivement. Les enseignants doivent en tenir compte afin de satisfaire à la fois aux exigences d'un cursus scolaire prédéterminé et aux besoins des différents élèves qui composent la classe, aux connaissances, antécédents et intérêts variés. Nous insistons ici sur ces deux conceptions opposées de la compétence, notamment sur leurs répercussions connexes sur le plan de l'épistémologie, de la pédagogie, de la psychologie et de l'évaluation. Cette démarche démontre que l'évaluation en classe est un véritable paradoxe: elle cherche à atteindre des objectifs conflictuels qui sont intimement liés aux différentes conceptions de la compétence.

Marks really do not fit with the thrust of the curriculum we are using.

How do you know that a kid "appreciates?" That is really, really tough to measure.

The interim report asks for satisfactory or not satisfactory. Well, what is satisfactory? Who defines what is satisfactory? Is it 50s, 60s, 70s? It's not really clear. And if it's not clear to us it's not clear for the parents. . .

I modified the program so they can be successful. Well, what is successful? Is successful 50%, 60%, 80%? When you modify for those children so that they are successful, if you have a watered down program to the extent that these children are, on paper, a success, and they go off to high school and those modifications aren't met, you've watered it down to the extent that it's not a true outcome. (from Earl & Katz, 2000)

While assessment of students in terms of expressed goals has always been a central feature of schooling, recent concern with accountability has led to an increased focus on programs that assess students in universalized terms against “objective” standards. The outcomes of such assessment programs could then be used to judge not only children relative to each other, but also classrooms, schools, provinces, and nations. Teachers, understandably, are poised halfway between the children they serve and the mandated “objective” standards. The result is the kind of conflict cited in the epigrams given above. In this paper, we identify the roots of this conflict by focusing on the assessment arena in which teachers are most directly implicated – classroom assessment. Specifically, we will argue that the problematic nature of classroom assessment stems from its attempt to serve competing purposes which are intimately bound to alternative conceptions of competence.

It is in the preservation and accumulation of competencies that culture is made possible (Premack & Premack, 1996; Tomasello, Kruger & Ratner, 1993). Schools are the institutions charged with the transmission of these competencies from generation to generation. Conceptions of competence have changed with advances in educational and psychological theory. There are at least two very different conceptions of competence at play on the landscape of formal education, best defined in terms of what is known collectively and what is known subjectively. Teachers are forced to attend to each of these in their challenge to meet both the demands of a fixed school curriculum and the needs of the individual children with their varied understandings, backgrounds, and interests who make up a class (Olson & Katz, 2000). In the pages that follow, we outline these two contrasting notions of competence, including their associated epistemological, pedagogical, psychological, and assessment implications. In this way, the paradox of classroom assessment is exposed. We begin with “the known.”

The known

Now, what I want is Facts. Teach these boys and girls nothing but Facts. Facts alone are wanted in life. Plant nothing else, and root out everything else. You can only form the minds of reasoning animals upon Facts: nothing else will ever be of any service to them. . . . (Dickens, 1854, p. 11)

So went the opening rant of Dickens' character, Mr. Thomas Gradgrind, as he presented himself to “. . .the little pitchers before him, who were to be

filled so full of facts." Although, these days, no one urges just the facts, competence defined in terms of the possession of knowledge is a well-established feature of modern day schooling. Students arrive at school with minds considered devoid of the necessary facts, rules, and principles which must then be transferred in a unidirectional fashion from the instructor, text, or other authority. Teaching becomes an exercise in telling, and learning an exercise in remembering (Baxter Magolda, 1992; Olson and Katz, 2000). The sequential and hierarchical structure of didactic teaching exemplifies the pedagogy at work here, an image well captured in Bernstein's (1972) "collection codes" view of knowledge. The "collection codes" approach emphasizes the acquisition of hierarchical sequences of information specific to given disciplines.

The substance of what is told (taught) and what is remembered comes in the form of propositional knowledge, knowledge that is taken to have an existence independent of the individual knower. The route to independence is one that has received much attention in the postmodern literature. Berger and Luckmann (1967) highlight the dialectic relation that exists between "objective facticity" and "subjective meanings" where the former, which come to be seen as natural and unquestionable, are in fact conventions that originate as the latter and come to be solidified by various means. In much the same way, sociologists of science Bruno Latour and Stephen Woolgar (1986) trace the arduous journey from private conjectures such as "I wonder if. . ." to canonical truths, "The properties were shown to be. . ." The final report, as Smith (1990) points out and as noted above, removes all traces of subjectivity to become statements of fact, constituting "the known."

Despite such illuminating work from sociologists of knowledge, both schooled and unschooled arenas continue to rely on the assumption that knowledge is found not made, that truth is "the truth," not that which is taken to be true. Latour and Woolgar (p. 240) have concluded that ". . . the result of the construction of fact is that it appears unconstructed by anyone," and this observation is borne out by the extensive epistemological belief literature which suggests that few people ever doubt this traditional view of "the known" (Baxter Magolda, 1992; Belenky, Clinchy, Goldberger, & Tarule, 1986; Kitchener, 1983; Kitchener & King, 1981; Kitchener, King, Wood, & Davidson, 1989; Kuhn, 1991, 1992; Perry, 1970). For most, knowledge is certain and permanent (Schommer, 1994), and viewed as either right or wrong. Chandler, Boyes, and Ball (1990, p. 377) caricature this "absolutist" statement of knowledge as a ". . . free-standing attribute of the enviroing world that only secondarily comes into the passive possession of those who, because they happen to be in the right place at the right time, automatically end up with some portion of the unmitigated truth directly embossed upon the recording equipment of their minds." Disagreements between people, then, can be attributed to the lack of facts among members of one of the

parties; the reins of certainty can be pulled in by reference to a disagreeable position as “just your opinion.”

The process by which knowledge is authenticated as “the truth” in this view is thought to lie with the appropriate authorities (Baxter Magolda, 1992; Kitchener, 1983; Kitchener & King, 1981; Kitchener et al., 1989; Kuhn, 1992; Perry, 1970). Belenky et al. (1986) call this “silent knowing,” in which the individual accepts the authority’s proclamation as to what is true. There is no belief that the knower can learn from his or her own experiences and knowing does not belong to the individual. Thus, through an external definition of “the known,” clear specifications as to what is to be taught appear. Nowhere is this manifestation more evident than in the establishment of a formal curriculum embodying those truths.

What is especially interesting from our assessment-centred perspective is that in providing such clear specifications of “the known” as that which must be taught, the formal curriculum also provides an associated set of standards for assessing learning achievements. That is, classroom assessment becomes the vehicle for gauging the closeness of fit between what is told and what is remembered (Earl & Katz, 2000). In such contexts, even the evidence used to justify beliefs as “knowledge” is traceable to authorities. Since “the truth” is acquired from appropriate authorities, justification for belief normally takes the form of what Belenky et al. (1986) have termed “received knowing.” That is, evidence consists of congruity with the authority, be it person or text.

While it is true that formal testing is the typical manifestation of the “concordance check,” a similar check is built into the anecdotal assessments that stem from classroom discourse. A common classroom discourse pattern is the IRE sequence (initiate, respond, and evaluate), where the teacher asks a question, a student responds, and the teacher gives feedback (Alvermann, O’Brein, & Dillon, 1990; Cazden, 1988). Unlike normal questions which are honest requests for information, in classroom questioning teachers are aware of the answers they are looking for (Applebee, 1996). Thus, Gradgrind’s request for his star, “well-crammed” pupil, Bitzer, to define a horse has nothing do with Grandgrind’s knowledge, or lack thereof, in the equine domain and everything to do with gauging the congruity of Bitzer’s definition of a horse with that given by the dictionary as:

Quadruped. Graminivorous. Forty teeth, namely, twenty-four grinders, four eye-teeth, and twelve incisive. Sheds coat in the spring; in marshy countries, sheds hoofs, too. Hoofs hard, but requiring to be shod with iron. Age known by marks in mouth. (Dickens, 1854, p. 14)

The point is further illustrated by the experiences of those who have worked with younger children and received one of the following answers to a less than honest “why” question: “Because you said so” or “Because the book said

so.” Finally, assessment as judgment, particularly as supported by classroom questioning, is a culture specific phenomenon. Heath (1982, p. 105) offered the following transcript as evidence for African-American students’ expectation that questions have answers unknown to the questioner:

Teacher: What is the story about?

Children: (silence)

Teacher: Uh... Let’s see... Who is it the story talks about?

Children: (silence)

Teacher: Who is the main character: Um... What kind of story is it?

Child: Ain’t nobody can talk about things being about themselves!

Such failed questioning series were also reported by Luria (1976) in his studies on non-literate adults.

By defining that which is “known” and subsequently what must be taught, modes of education predicated on the formal curriculum are essentially selective (Glaser & Silver, 1994). That is, there is minimal variation in the conditions of learning and consequently, in instruction. We have, as just one illustration, the provincial government of Ontario, Canada introducing a highly centralized, domain-specific provincial curriculum which spells out exactly what students will know and when. For example, in mathematics, by the end of Grade 8 students will “represent whole numbers in expanded form using powers and scientific notation (e.g., $347 = 3 \times 10^2 + 4 \times 10 + 7$, $356 = 3.56 \times 10^2$)” (Ontario Ministry of Education and Training, 1997, p. 26). The accompanying justification-driven media campaign explains the ease with which children will be able to move between school board jurisdictions given the new-found consistency in what is to be taught and learned.

Although pedagogically naïve, the virtue of such a specific curriculum is its direct relationship to what must be taught. By specifying this precisely, it sets the stage for a rather specific definition of accountability, accountability defined as a kind of surveillance mechanism which proceeds by way of a congruity check. What is prescribed by the formal curriculum necessarily maintains an existence apart from individual teachers and their students. It is a prescriptive blueprint for common action and by definition a yardstick for common evaluation, hence systemic accountability. In this way, it is possible to begin to talk about a target competence “norm” or “benchmark,” defined not in terms of characteristic variability, but rather in terms of central tendency – the “mean” not the “spread” of the distribution, so to speak. In addition to the classroom routines we have described, such thinking is responsible for the establishment and maintenance of the large-scale achievement testing movement (c.f. Glaser & Silver, 1994). Assessment is about judgment, judgment about who will go on and where they will go, or, to put it differently, about who is competent and who is not.

While it is true that the ideal manifestation of the achievement of fixed knowledge would appear in student performances characterized by predetermined means and zero variances (Hacking, 1996), diversity amongst learners in their achievement of these fixed goals is conspicuous. Here enters the appeal of psychological theory; in particular, the psychology based on intelligence and personality testing that serves primarily to predict and explain away the differential outcomes of schooling. We are referring here to that stance to the problem of diversity known as Individual Difference. Individual Difference proceeds by classifying people on the basis of values on particular dimensions or traits. It was trait ascription that allowed Binet and his successors to explain 25 percent of the variance in “typical” school-like tasks by reference to “intelligence,” and it is trait ascription that characterizes much of teacher talk today as they attempt to account for classroom performance variance on the basis of particular dimensions like “persistence,” “impulsivity,” “giftedness,” or “hyperactivity” (Olson, 2000). Although such characterization is almost irresistible and often harmless, it is also misleading, especially when such traits are concretized into entities. Intelligence, for example, began as an abstract relation between strategies and tasks but quickly became a possession, or perhaps more accurately, a capacity with implications for competence. In doing so, it proffered an explanation for variability, while allowing the tradition of achieving a fixed norm to continue.

The picture of formal education we have provided thus far is one familiar to anyone who has read critiques of schooling from Dewey to the present. It is one circumscribed in a notion of competence defined in terms of an individual’s acquisition of knowledge construed as facts and theories, i.e. “the known.” Attempts to define “the known” result in clear specifications as to what is to be taught, as well as an associated set of standards for assessing learning achievements. Gauging the fit between what is taught and what is learned defines the nature of classroom assessment, a process which serves as a mechanism for making competence judgments. Diversity, in this framework, is a departure from the ideal of a “norm” and requires explanation, explanation that primarily appeals to trait psychology – the psychology of happenings (Olson, 2000).

While today’s classrooms exemplify the educational vision detailed above, they are also characterized by more than simply a concern for “the known.” Indeed, such a concern results in subjective postures sharply distinguished from the objectivist stance which relegates personal belief to the substandard position of opinion, or worse, idle speculation. Thus, young Sissy Jupe’s inclusion of her experiences with the animal in her definition of a horse – in particular her father’s “belonging to the horseriding” – meets with the Gradgrind reply, “Girl number twenty unable to define a horse! Girl number twenty possessed of no facts, in reference to one of the commonest of

animals!” (Dickens, 1854, p. 13). However, much of what we see in today’s classrooms accepts, indeed encourages, what is known subjectively and experientially and seeks to develop it. We detail next this concern for “the knower,” noting some very different epistemological, pedagogical, psychological, and assessment definitions from those articulated above.

The knower

The image of respect for the child’s point of view is one well captured in the writings of John Dewey. Influenced by Darwin’s theory of evolution and the psychology of William James, Dewey spoke of thinking and learning as practical capacities through which we actively interact with our surroundings (Phillips & Soltis, 1991). In nature, he argued, thinking is stimulated by real problems that the learner has a personal vested interest in solving (Dewey, 1902/1966).

Dewey argued that traditional schools restrained both mental and physical activity. He saw knowledge emerging from a process of interpretation and clarification of meanings related to various aspects of experience in the world (Dewey, 1938). Along with this emphasis on interpretive sense-making came a psychology of **doings** that stood in contrast to the psychology of **happenings** (Olson, 2000). Dewey argued that while trait psychology proffered causal mechanisms that explained behaviour in terms of what happened, a preferable alternative would be to explain behaviour in terms of what agents, including learners, were doing or trying to do. Thus, the shift is from causes to reasons, from persons as passive respondents to persons as intentional agents. Children are seen as acting in accordance with their beliefs, desires, hopes, and intentions, that is, their mental states.

Child-centredness is, by now, a dominant theme in both cognitive-developmental and sociocultural perspectives of education (Astington & Pelletier, 1996; Shuell, 1996). In the former, Piaget’s genetic epistemology comes close to providing a psychological theory that parallels Dewey’s educational one. Cognitive developmentalists emphasize the child’s interactions with the physical environment. Learning occurs as previously acquired cognitive structures are coordinated to form new superordinate structures. Initially, structures develop from concrete experiences and thinking remains concrete until the coordination of structures allows for the emergence of abstract thought. In the latter, the sociocultural perspective, the child’s interactions with other people in a social world are emphasized. Rooted in Vygotskian theory, cognitive development is regarded as a socially mediated process. The conversation metaphor (Applebee, 1996) is a popular one, with participation in classroom discourse serving as the vehicle through which meanings are constructed. Notions of cognitive apprenticeships (Collins, Brown, & Newman, 1989) and communities of learners (Brown & Campione, 1994; Rogoff, Matusov, & White, 1996) portray learning as the

result of a coordination of perspectives between teacher and learner or amongst learners themselves.

In considering the child's point of view, the mind is acknowledged as a place of privately held beliefs and ideas (Olson & Bruner, 1996; Olson & Katz, 2000). Children are seen as individuals capable of sense-making both through their interactions with the environment (cognitive-developmental) and through discourse with others (sociocultural). Learning, then, can be conceptualized as a process of subjective interpretation, while teaching creates opportunities for intersubjective interchanges, for sharing beliefs through collaborative discourse.

Defining competence in terms of the subjective properties of the knower leads to a very different set of epistemological assumptions than those discussed in the previous section. Indeed, in recognizing that "the known" is neither God-given nor an indisputable fact of nature, knowledge is viewed as fundamentally made, not found, and consequently uncertain. Truth is seen as varying from person to person and knowledge is interpreted subjectively (Kuhn, 1991, 1992; Perry, 1970). Constructivism, in this view, is the key notion.

From this relativistic epistemic stance, the source of knowledge lies in personal, idiosyncratic processes such as individual opinion (Baxter-Magolda, 1992; Kitchener, 1983; Kitchener & King, 1981; Kitchener et al., 1989; Kuhn, 1992; Perry, 1970). Belenky et al. (1986) use the term "subjective knowing" to convey the delimiting role of intuition and gut feeling in the epistemological enterprise. Personal belief counts as evidence for knowing. To justify beliefs is to share them with others. Competence is seen as subjective sense-making.

The subjectively-driven notion of competence provides a new stance to classroom assessment. First, it shifts the focus from desired mean to desired variance, a view of diversity that is supported by an intentional psychology. Diversity is a central feature of an intentional psychology. Deviations from a predetermined norm are not aberrations but rather expectations. Normalcy becomes a property of variability rather than a central tendency because learners as interpreters necessarily construct somewhat different interpretations of any text or event. Classroom assessment is the mechanism that promotes the interpretation and revision processes of the learner. It is something best accomplished by the self since it is the knower who holds privileged access to the relevant beliefs and intentions, though they are also obliged to give a public account (c.f. Lampert, Rittenhouse, & Crumbaugh, 1996). Thus, portfolios and other self-assessment vehicles (Earl & Cousins, 1996; Garcia & Pearson, 1994; Glaser & Silver, 1994; Wilson, 1996) which promote students' articulation of subjective reasons for their beliefs and opinions rather than objective right answers (Buchmann, 1986) are sug-

gested. Additional assessment opportunities can also be found in classroom discourse episodes that support subjective interpretation.

When subjective reasons replace objective right answers, assessment, as we have seen, becomes in large part the responsibility of the knower; the knower has to justify his or her views. Assessment opportunities can be found in classroom discourse patterns which change too, from IRE to ICE (Wilson, 1996) in keeping with acronym inscription. ICE discourse encourages students to form and share their own Ideas, Connect them to one another, and Extend them beyond personal experience (see Wilson, 1996 for an example). Gardner (1990) has proposed “individually configured excellence” as a title reflective of this adaptive (Glaser, 1977) model of education. Schooling works to identify and develop sources of competence in individual students. It is in such subjective competencies that individual backgrounds, interests, and understandings are taken into account.

Conceptions of education driven by a concern for “the knower” are not without limitations. The epistemic posture associated with a subjectively-driven notion of competence derogates truth to a transient quality. Truth is variant and certainty is elusive. In this way, the stance of “skeptic” is legitimated in which “. . . all claims are challenged, every heart-felt belief is held up to ridicule, and all action is seen as permanently premature” (Chandler et al., 1990, p. 379). The problem here is that in the absence of certainty, all beliefs appear equal. Neatby (1953) points out that it is this very phenomenon that supports mediocrity or, as she puts it, a system of education that offers “so little for the mind.” Without objective standards, and the associated norms or benchmarks, systemic accountability becomes problematic for both individuals and educational institutions. It is in terms of such standards that schools are held accountable for fulfilling an institutional mandate, a mandate generally advanced in terms of those very standards.

The paradox of classroom assessment

The practice of classroom assessment is one that can be conceptualized as necessarily consisting of multiple purposes, hence the potential for emergent paradoxes. Wilson (1996) explains that assessment must satisfy many goals such as providing feedback to students, offering diagnostic information for the teacher to use, providing summary data for record keeping, proffering evidence for reports, and directing efforts at curriculum revision. He underscores the inextricable link between assessment and pedagogy that we have established throughout this paper by referencing five different but necessary assessment roles which teachers must play.

First, in “teacher as mentor,” encounters with individual students help teachers make up their minds about each student’s progress and about what

will be necessary for growth to occur. Such assessments are often not noted, recorded, or even recognized. They form part of the social fabric of the class. The second assessment role, "teacher as guide," involves managing the progress of the class as a group. Markedly different from the practice of providing individual help to a single student, this role requires forming and acting on impressions of the class's progress. Third, employing numerical categories for the purpose of reporting student achievement delimits the role of "teacher as accountant." The task here is to provide a common basis of comparison so that each child appears to be treated as fairly as every other child. Indeed, this role is intimately bound to the fourth one, "teacher as reporter," in which the teacher is required to keep organized records in order to expedite the process of gathering, interpreting, and relaying information for outside communication. Finally, the "teacher as program director" is responsible for reflecting on the classroom program as a whole. This task requires drawing on periodic external assessments which specify the systemic goals that help set the overall "tone" for the classroom program.

The potential for conflict among the many assessment roles that Wilson (1996) spells out seems unavoidable. Translated into the broader language of the present argument, we see that a concern for both the "knower" and the "known" is an inherent property of the individual classroom. Teachers are pulled between the two sets of constraints in that classes are required to meet fixed and predetermined goals, yet it is individuals with varied backgrounds, interests, and understandings that learn (Clay, 1996). This, we have argued, is problematic because each element maps onto a different conception of competence which, in turn, is intimately bound to a highly specific epistemology, pedagogy, and psychology. Contradictions inherent in classroom assessment practices are unavoidable. They come to us by natural extension of incommensurable competence definitions. When the concern is with what is "known" generally, assessment is about correspondence to a norm or standard. When the concern is with the "knower," assessment is a subjective mechanism of interpretation.

For much of this century educational theory has attempted to come to grips with this gap between what is known generally and should be transmitted, and what is known subjectively and should be developed. Yet this effort typically consisted of little more than staking opposing claims on the epistemological battleground. Classical education theory considered knowledge to be an entity that exists in some tangible form capable of being transferred intact to children. Educational reformers, such as Dewey, took the other pole, insisting on the primacy of subjective and communal experience as the basis of all knowledge.

Katz (1998) illustrates these oppositional tendencies in the context of a heated debate presently reflected in the mathematics education literature.

Specially, a tension exists between providing opportunities for mathematical understanding on the one hand, and the need to teach widely used notations and algorithms on the other, with proponents of the former arguing that the latter has no place in the classroom. However, the advance of mathematics is, in part, the invention of notations and algorithms that simplify our tasks, hence the need for students to learn them. Thus, although theoreticians have been more or less content to operate almost exclusively in spheres defined by one epistemology or the other, the practical landscape of education does not enjoy the bounded consistency of abstract theory. Both elements of the epistemological enterprise seem valid in a context of actual practice, hence the need for an alternative solution to the epistemological, and thus assessment, paradox. After all, assessment practices of all varieties communicate something about the learner to the learner. They each communicate a way of thinking about the self. They are the vehicles through which children come to think of themselves as competent on the one hand or as incompetent on the other. We offer the paradox described here as a challenge to educational researchers and practitioners in twenty-first century.

REFERENCES

- Alvermann, D., O'Brien, D., & Dillon, D. (1990). What teachers do when they say they're having discussions of content area reading assignments: A qualitative analysis. *Reading Research Quarterly*, 25, 296-322.
- Applebee, A. N. (1996). *Curriculum as conversation: Transforming traditions of teaching and learning*. Chicago: The University of Chicago Press.
- Astington, J. W., & Pelletier, J. (1996). The language of the mind: Its role in teaching and learning. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 593-619). Cambridge: Blackwell.
- Baxter Magolda, M. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey Bass.
- Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.
- Berger, P., & Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. New York: Anchor.
- Bernstein, B. (1972). On classification and framing of educational knowledge. In M. Young (Ed.), *Knowledge and control* (pp. 47-69). London: Collier/Macmillan.
- Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-270). Cambridge: Bradford Books/MIT press.
- Buchmann, M. (1984). The use of research knowledge in teacher education and teaching. *American Journal of Education*, 92, 421-439.
- Cazden, C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann.
- Chandler, M., Boyes, M., & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology*, 50, 370-395.

- Clay, M. (1996). Accommodating diversity in early literacy learning. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 202-224). Cambridge: Blackwell.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum.
- Dewey, J. (1902/1966). *The child and the curriculum*. Chicago: University of Chicago
- Dewey, J. (1938). *Experience and education*. New York: Collier Books.
- Dickens, C. (1854). *Hard Times*. New York: Penguin.
- Earl, L., & Cousins, J. B. (1996). *Classroom assessment: Changing the face; facing the change*. Toronto: Ontario Public School Teachers' Federation.
- Earl, L., & Katz, S. (2000). Classroom assessment: Teachers' struggles to change. In N. Bascia & A. Hargreaves (Eds.), *The sharp edge of change*. London: Falmer.
- Garcia, G., & Pearson, P. (1994). Assessment and diversity. *Review of Research in Education*, 20, 337-391.
- Gardner, H. (1990). The difficulties of school: Probable causes, possible cures. *Daedalus: Journal of the American Academy of Arts and Sciences*, 119, 85-113.
- Glaser, R. (1977). *Adaptive education: Individual diversity and learning*. New York: Holt, Rinehart & Winston.
- Glaser, R., & Silver, E. (1994). Assessment, testing, and instruction: Retrospect and prospect. *Review of Research in Education*, 20, 393-419.
- Hacking, I. (1996). Normal people. In D. Olson, & N. Torrance (Eds.), *Modes of thought*. Cambridge, UK: Cambridge University Press.
- Heath, S. B. (1982). Questioning at school and at home: A comparative study. In G. Spindler (Ed.), *Doing the ethnography of schooling: Educational anthropology in action* (pp. 102-131). New York: Holt, Rinehart & Winston.
- Katz, S. (1998). Substituting the symbol for the experience: Exposing a fallacy in mathematics education. *Journal of Mathematical Behavior*, 17, 405-410.
- Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three level model of cognitive processing. *Human Development*, 4, 222-232.
- Kitchener, K. S., & King, P. A. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology*, 2, 89-116.
- Kitchener, K. S., King, P. A., Wood, P. A., & Davidson, M. L. (1989). Sequentiality and consistency in development of reflective judgment: A six-year longitudinal study. *Journal of Applied Developmental Psychology*, 10, 73-95.
- Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press.
- Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, 62, 155-178.
- Lampert, M., Rittenhouse, P., & Crumbaugh, C. (1996). Agreeing to disagree: Developing sociable mathematical discourse. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 731-764). Cambridge: Blackwell.
- Latour, B., & Woolgar, S. (1986). *Laboratory life: The social construction of scientific facts*. Princeton: Princeton University Press.
- Luria, A. (1976). *Cognitive Development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.
- Neatby, H. (1953). *So little for the mind*. Toronto: Clarke, Irwin.
- Olson, D. R. (2000). There are x kinds of learners in a single class: Diversity without individual differences. In J. Gaffney & B. Askew (Eds.), *Stirring the waters: A tribute to Marie Clay*. Portsmouth, NH: Heinemann.

The Paradox of Classroom Assessment

- Olson, D. R., & Bruner, J. S. (1996). Folk psychology and folk pedagogy. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 9-27). Cambridge: Blackwell.
- Olson, D. R., & Katz, S. (2000). The fourth folk pedagogy. In B. Torff & R. J. Sternberg (Eds.), *Understanding and teaching the intuitive mind*. Hillsdale, NJ: Erlbaum.
- Ontario Ministry of Education and Training. (1997). *The Ontario Curriculum Grades 1-8: Mathematics*. Toronto: Queen's Printer.
- Perry, W. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart & Winston.
- Phillips, D., & Soltis, J. (1991). *Perspectives on learning (2nd edition)*. New York: Teachers College Press.
- Premack, D., & Premack, A. J. (1996). Why animals lack pedagogy and some cultures have more of it than others. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 302-323). Cambridge: Blackwell.
- Rogoff, B., Matusov, E., & White, C. (1996). Models of teaching and learning: Participation in a community of learners. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development* (pp. 388-414). Cambridge: Blackwell.
- Schommer, M. (1994). Synthesizing epistemological belief research: Tentative understandings and provocative confusions. *Educational Psychology Review*, 6, 293-320.
- Shuell, T. (1996). Teaching and learning in a classroom context. In D. Berliner & R. Calfee (Eds.), *The handbook of educational psychology* (pp. 726-764). New York: Macmillan.
- Smith, D. (1990). *Conceptual practices of power*. Toronto: University of Toronto Press.
- Thomas, G. (1992). *Effective classroom teamwork: Support or inclusion?* London: Routledge.
- Tomasello, M., Kruger, A. C., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495-511.
- Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership*, 49, 26-33.
- Wilson, R. (1996). *Assessing students in classrooms and schools*. Toronto: Allyn & Bacon.

STEVEN KATZ is a doctoral candidate in the department of Human Development and Applied Psychology, and a member of the International Centre for Educational Change, at the Ontario Institute for Studies in Education of the University of Toronto. His research focuses on the epistemological foundations of teachers' instruction and assessment practices, and proceeds in contexts delimited by broader policy initiatives.

LORNA EARL is an Associate Professor in the department of Theory and Policy Studies and a Director of the International Centre for Educational Change at the Ontario Institute for Studies in Education of the University of Toronto. Her research interests are in the wise application of research, assessment, and evaluation knowledge to the realities of schools and classrooms.

DAVID OLSON is a Professor in the department of Human Development and Applied Psychology at the Ontario Institute for Studies in Education of the University of Toronto. His research is concerned with literacy and the development of children's theories of mind.

STEVEN KATZ est étudiant au doctorat au Département du développement humain et de psychologie appliquée, et membre du Centre international de changement pédagogique à l'Institut d'études pédagogiques de l'Ontario de l'Université de Toronto. Ses recherches portent sur les fondements épistémologiques des méthodes d'enseignement et d'évaluation, celles-ci étant resituées dans le cadre plus général des politiques et initiatives stratégiques affectant l'éducation.

LORNA EARL est professeure agrégée au Département d'études théoriques et politiques et directrice du Centre international de changement pédagogique à l'Institut d'études pédagogiques de l'Ontario de l'Université de Toronto. Ses recherches portent sur l'application raisonnée des connaissances en matière de recherche, d'évaluation et de mesure des apprentissages aux réalités des écoles et des salles de classe.

DAVID OLSON est professeur au Département du développement humain et de psychologie appliquée à l'Institut d'études pédagogiques de l'Ontario de l'Université de Toronto. Ses recherches portent sur la littératie et l'acquisition des théories mentales chez les enfants.