Accountability to Whom?  
Testing and Social Justice  
A Response to “Imagining No Child Left Behind Freed from Neoliberal Hijackers”  

David W. Kritt

Abstract
In response to Eugene Matusov’s article in this journal, Kritt addresses assumptions of the large-scale testing central to NCLB. Discussion of studies of urban kindergarten children that examine cognitive variability, including the assertion of ability, focuses on how this affects the student as a learner, as well as a teacher. In contrast, Matusov questions root assumptions of schooling, casting engagement in socially valued activities as an issue of human rights. This view is criticized as overly socialized. It is argued that surface-level functioning in a cultural context is not sufficient for full participation in a democracy.

The acrid rhetoric of educational politics certainly did not begin with No Child Left Behind (NCLB). But the Bush administration’s provocation in introducing NCLB—citing “the soft bigotry of low expectations” (U.S. Department of Education, 2003)—shifted the dialogue about educational policy to cast those currently teaching, and especially their unions, as the primary enemy. Eugene Matusov (2011) takes the 2001 NCLB law at its face value, as the place to “start a discussion about an alternative, genuine, sociocultural vision of failure-free education” (p. 7). My engagement in that discussion is grounded in practices in urban schools, the gap between how children actually learn and what tests measure, and a radically different interpretation of what tests reveal. Matusov and I agree that testing fails to capture thinking at its best, and it certainly should not be guiding instruction. And we generally share a sociocultural perspective. But I believe he presents an overly socialized viewpoint and so, we arrive at somewhat different conclusions.

A brief history reveals some of the assumptions of large-scale testing. The original impetus for intelligence tests was educational placement of what, today, we often call special-ed students. The initial widespread use of intelligence tests was in the military during World War I, to determine who should be sent for specialized training, for example, in communications, leaving those not selected to do more routine and mundane tasks. In its long and ignoble history, so-called intelligence testing has also been used as an exclusionary tool to restrict immigration and eligibility to vote and as a justification for proposed eugenics programs. In these cases, the negative implications of testing for democracy can be clearly seen. In educational practice, this is less evident but of no less serious concern.

The logic of testing has been to try to pass everyone through the same sieve. All individuals are expected to answer the same set of questions to demonstrate proficiency at the same mental tasks. This is rife with assumptions of homogeneity, both in the population and in the requirements of jobs and careers. In educational settings, intelligence testing and achievement testing have been used for sorting students into ability groups. Some students are identified as having deficits requiring placement in special-education classrooms. Others are tracked into classes with students of similar proficiency levels in reading, math, or science.

Matusov questions two common statistical gambits used in reporting test results. One is normalization. A common standard is set for all children, but criteria for passing may be adjusted to
ensure that the range of performance fits a normal curve, with some students surpassing the mean and some students failing. Although the tails of the distribution contain relatively few individuals, the competitive assumptions are anathema to Matusov. In a similar vein, he objects to the goal of closing achievement gaps, because it implicitly accepts proportional success/failure among social groups (e.g., Black-White, male-female, rich-poor); in other words, some within each group are expected to fall below the mean.

In several studies of urban kindergarten children (Kritt, 2004; Kritt, under review; Kritt & Shulman, in preparation), I examined cognitive variability in a more fine-grained manner. Using disaggregated testing data to look at individual performance (i.e., a collection of quantitative case studies) allowed for interrogation of the assumption that there is a relatively homogeneous mass of students with moderate abilities. A central debate regarding intelligence has focused on whether it is a general factor, so that we'd expect a person to have many talents in seemingly unrelated areas, or one that is domain specific. In the past, it was more common among psychologists to view a person as generally intelligent or not (Buckhalt, 2002; Carroll, 1991; Humphreys, 1979; Kranzler & Jensen, 1991). This idea has stuck among the general public, and may be at least partially due to a bias in social cognition, so if a person is good in one thing we simply assume he or she is good in many things. For example, Einstein was also, reportedly, a pretty good violinist, and comedian Steve Martin is also an actor, author, and musician. However, psychologists have increasingly recognized that many talents are domain specific (Gardner, 1983; Ceci, 1990). This may be due to either nature or nurture or, most likely, some combination of the two.

The findings of the aforementioned studies, that there is a great deal of variability in an individual's test performance across several areas (verbal, mathematical, perceptual-performance, memory), support an assertion by Gould (1981) that the statistical ideal of the average student, with abilities near the mean in all skill areas, is not representative of actual individuals. It is an important point because increasing numbers of students have difficulties in school that are subsequently labeled (e.g., as learning, attention, and behavior disorders). Now, all this diagnosis of children might be due to greatly improved diagnostic techniques. But Gould suggested in a later book (1996) that children never were uniform in their abilities. And Howard Gardner (1991) argued that schools have failed to adequately educate the vast majority of students.

Several implications for educational practice are rather obvious. Foremost, teachers should recognize a wide range of individual proficiencies and goals and facilitate their development. Progressive educators encourage students learning while being involved in rich inquiry (e.g., projects), as well as emphasize authentic assessment that is part of meaningful classroom activities (e.g., Allen, 1998; Camp, 1990; Grigorenko & Sternberg, 1998; Kritt, 1993; Paris et al., 1992; Wiggins, 1989, 1998; Wolf, 1989). It must be recognized that testing is an imposed and decontextualized task with no intrinsic meaning to the test taker, only instrumental value, as a means to some rather abstract academic-achievement goal. Some students may embrace this imposed goal as part of an achievement ideology (i.e., belief that doing well in school leads directly to doing well in life), but many, especially in urban populations, reject it (e.g., MacLeod, 1987/2008).

Rather than summative assessments that serve actuarial functions for politicians, assessment can be formative, so that it contributes to student learning and development. Vygotsky's (1978) zone of proximal development (ZPD) is a prime example of assessment that could be useful to adults (parents and teachers) truly concerned with student learning (Kozulin, 2005; Rogoff, 1990; Wertsch, 1985). Instead of stopping the assessment with a child's answer to a question, the process continues to determine what the child can do with cuing and gentle assistance. The distance between what the child can do alone and what can be done with assistance is the ZPD. Educational efforts (e.g., scaffolding) are targeted here.

Consideration of some classic issues in intelligence help us to further understand the way students are characterized and educated. One is whether intelligence is inherited or dependent on the environment. A famous family, the Huxleys, illustrates why it is a thorny issue. The grandfather was a biologist and an early defender of Charles Darwin's theory of evolution. His grandsons included: Sir Julian Huxley, a noted biologist; Aldous, the author of Brave New World, The Doors of Perception, and other novels; and little brother Andrew, who received the Nobel Prize in physiology and medicine for his work on activity in the central nervous system. I think it is safe to say they had good genes, but also they undoubtedly benefited greatly from growing up in a stimulating family, being brought up in a way that encouraged curiosity and creativity. It is virtually impossible to separate genetic and environmental influences, with notable exceptions being studies of identical twins separated at birth. Focusing primarily on genetic inheritance ratifies not only privilege but also the futility of educating problem students. The faulty logic of arguments such as Herrnstein & Murray's (1994), that the current distribution of performance on IQ tests is indicative of hereditary differences between races in intelligence, serves to justify existing economic hierarchies (Fischer, Janowski, Lucas, Swidler, & Voss, 1998; Jacoby & Glauberman, 1995). Furthermore, it suggests that patterns of placement in schools, employment, status, and power reflect inherent differences in brains, rather than socially constructed and maintained divisions.

Another classic issue regarding intelligence is perhaps even more to the point. The assertion of ability refers to a hypothetical construct used to justify expectations. It is a useful concept if we want to talk about prediction of future performance in a variety of areas, ranging from educational achievement to employment and career to general well-being (cf. Sternberg, 1997). But that quickly becomes problematic (Riegel, 1976). As educators, we have a responsibility to provide optimal conditions for children to develop their talents. Yet we can observe and measure only what children have actually done. These instances of performance may, of course, differ according to contextual factors, testing conditions, and emotional factors. Even if a child is repeatedly tested under less than optimal conditions, and repeatedly performs poorly, it may be inaccurate to assert that the child lacks ability. Rather, we must
acknowledge that the child cannot perform under certain conditions. And it is important to remember that the acontextual presentation of problems in a testing situation stands in stark contrast to the highly contextual character of activities that individuals perform in the real world and find meaningful.

What does it mean to say someone has potential? The person has never done it, whatever “it” is, yet we state that he or she could do it. That is not exactly the same as saying we expect that child to do it. Parents have dreams for their children. Affluent parents can more easily do more than poor parents to give their children second and third chances. It is not just a curious coincidence that both test scores and academic performance have repeatedly been shown to correlate with parental income. From a social Darwinist perspective, this is readily accepted as confirming the logic of the hierarchical structure of society. Those of a less severe bent are likely to seek an explanation in the effects of poverty.

If students from more affluent families are viewed in terms of their potential and other children viewed only in terms of their actual performance, they will be treated differently in educational settings. Teacher expectations can lead to favored treatment in the classroom that, in turn, contributes to student academic achievement (Alvidrez & Weinstein, 1999; Entwistle & Alexander, 1993). Differences between which students are most frequently called upon in class and which are not and among types of feedback given, as well as possible bias in appraisal of student work are mundane and often unintentional ways in which this occurs. Yet such microlevel dialogic processes contribute to macrolevel processes whereby schools are trading posts where economic and social privilege are transformed into superior academic performance (Bourdieu, 1977), which is used to justify superior life opportunities (cf. Anyon, 1981).

This same phenomenon can also be considered in a very different way. The use of IQ test results to predict future development is predicated upon implicit assumptions (e.g., emphasis on inheritance of intelligence, the g factor) about continuity of ability and the self. The effect of such beliefs on how individuals think of themselves as learners can influence students, parents, and teachers in ways that can accentuate the possibility of change or erect impediments to it.

We tend to view ourselves as others view us (Blumer, 1969; Mead, 1934/1962). This happens as a natural part of social interaction. If someone treats us as if we are smart and capable, we tend to give an optimal performance. If we are treated as if we are dumb, we tend to perform at a lower level than we might otherwise (e.g., Steele, 1997). This is not a necessary reaction, and some very resilient individuals can defy expectations, but those are more the exception than the rule.

In everyday interactions in most of life’s spheres, the judgment of others is a subtler process than it is in school. The apotheosis of being judged is formal labeling in schools. Today it is usually not blatant, but both labeled individuals and peers recognize it nonetheless. The labeled individual tends to behave in ways that confirm the label. That is not because the label is accurate but because a social dialectic has been set in motion. Adults let a student know they think the student can’t do something. Either out of defiance or out of laziness, that student does not do much. That is taken as further evidence, and maybe teachers will give the student easier problems, which are not successfully completed once again. And so a transactional cycle is set in motion. Such insights inform several major theories of motivation.

These theories have clear implications for understanding some extremes of student behavior in the classroom. Weiner (1979; 1986) focused on how people think about their individual successes and failures, whether they attribute their success to stable aspects of themselves, to stable aspects of the outside world, or to transient aspects either internal or external. Dweck (1986) identified individuals who think their own abilities are relatively fixed and others who view their performance as something that can change. Only the latter individuals consider the time and effort required to do homework to be worthwhile. Covington (1985) viewed it as the student’s attempt to protect a sense of self. If intelligence is an unchanging attribute, then students will use various tactics to avoid looking like they’re trying, so they will not appear, especially to peers, as incapable.

Reified ideas about a person’s own intelligence are largely unwarranted but are frequently established early in an individual’s school career (e.g., Bandura, 1993; Entwistle & Alexander, 1993). Such static conceptions can be detrimental to future learning and development. Even in the absence of tracking, such designations affect individuals’ concept of themselves as thinkers and learners, as well as how others (parents, teachers, peers, coworkers) interact with them. Students who have more fluid ideas about thinking and learning are more likely to try harder, explore, and be creative.

Behavior is malleable, performance is situation specific, and an individual’s functioning is best understood in its cultural context (and sometimes more than one, as with the embedded and interacting contexts of family, school, community, and macroeconomic forces). Carrying these insights to their logical conclusions, Matusov (2011) states that it is desirable to create a world where there is equality of “access to socially valuable practices” (p. 4). Following critics of the logic of disability, such as McDermott and Varenne (McDermott, 1993; McDermott & Varenne, 1995; Varenne & McDermott, 1998), Matusov considers this “an issue of human rights rather than one of education” (p. 4). And this bold stance certainly changes the terms of the discussion. By shifting the focus to “supportive infrastructures” and “policies and practices” and calling for “distinctive human networks” (cf. Hutchins, 1991; Lave & Wenger, 1991; Rogoff & Lave, 1984; Salomon, 1993), Matusov establishes that “learning is only one of many means for access” (p. 4). Citing aspects of the learning process that separate it from functional daily activity and the importance of communal rather than individual endeavor, he questions root assumptions of schooling as we commonly know them.

This focus on an ecology of support is laudable, but problematic in several respects. Matusov’s idealistic assumption of widespread support for failure-free education flies in the face of the very neoliberal values he critiques. In short, it seems apparent there is not enough to go around, and most parents want to give their own children an advantage. Recognizing sources of fundamental unfairness in capitalist society does not mitigate the deep influences
upon individual psychology (e.g., self-interest, competitiveness). There is the world we live in and the world we’d like to live in.

In the process of making his point, Matusov discusses a number of ideas very familiar to progressive educators (e.g., Duckworth, 1987/2006; Langer, 1997; Wells, 1999), including the need for meaningful context and tasks as well as student involvement in setting goals. He is also perhaps a bit politically incorrect, but truthful, when he states that “there will always be students who might not learn” (p. 6). Matusov sets himself apart when he states that “for these students, not to be behind is to develop other ways of access to socially and personally valuable practices and activities” (p. 6). For example, he suggests that reading is “one of many possible ways to access printed texts.” A glib counterexample is that I had to reread sections of his paper several times to grasp certain points. Hypothetically, I could have had access to this paper in aural form, but I am unsure if that would have allowed me to critically analyze the argument as well as I could using traditional literacy skills. Similarly, the use of some combination of images and enactment may have a more powerful impact but afford less opportunity for reflection and assessment of the message. Indeed, because modes of communication made possible by information and communication technologies are so immediate, it is more difficult to hold their messages at arm’s length and peruse them. Accordingly, some measures that might ensure education is failure free might also remove thought from the enterprise of engaging in socially valued activities.

Although ostensibly the same results might be achieved in some easier way, there is a difference between a surface-level performance and a deep and flexible understanding. Reading a plot synopsis is not equivalent to reading great literature. Being able to estimate the total for items scanned in a checkout line is a useful skill. Knowing enough history to evaluate pronouncements about the war in Afghanistan adds valuable perspective to simply seeing a report on CNN. And having the critical skills to compare the speeches of politicians makes a person a better informed voter (i.e., “having all the input” is not sufficient).

At best, as in pursuit of a hobby, learning is largely intrinsically motivated and self-directed. But students must demonstrate competence in a range of domains in order to graduate from high school and pursue advanced training or education. Those who favor a standardized education, hoping it will lead to very similar outcomes, will inevitably be disappointed. For even with comparable instruction and a level playing field, such an education is likely to have differing effects. What students can make of their education is dependent upon the prior experiences and frameworks they bring to it. A chess master viewing a chess match, an accomplished musician at the symphony, a chef eating a meal—all can find significance (and learn) in ways that the rest of us cannot. Even presented with identical situations, everyone does not benefit in identical ways. There can be multiple outcomes and there can be more than one pathway to a common outcome (cf. Werner, 1926/1948).

Fully aware of all the constraints and injustices of the world, as an educator and as a teacher of teachers, I have a primary objective: I want young people to have options in life. This does not always require merely academic skills. Nor, as Matusov suggests, is it simply being able to do what everyone else can (cf. a Turing test for the illiterate or innumerate). It certainly does require a well-developed facility in questioning what is presented and thinking outside the box when necessary. Access and equality may be steps in this direction, but they are not sufficient preparation for full participation in a democracy.

Matusov’s analysis is ultimately based upon material dialectics. In these times of emphasis on market forces in education (e.g., charter schools), this is a fresh and provocative perspective. Yet a more cognitive perspective upon cultural-historical activity theory (i.e., informed by Vygotsky’s writing on scientific concepts, 1934/1987, as well as by Friere, 1970/1986) suggests that psychological change is not automatic and contingent upon material conditions. Rather, generalization, systematization, and understanding of abstract principles are necessary for people to become truly conscious, in the sense of being able to reflect upon their own behavior and thought. In addition to preserving some of the best values of the educational tradition, this sort of awareness is necessary for the kind of real choices required for participation in a democracy. When we work toward creating the conditions for true and deep understanding that is not divorced from functioning within the cultural contexts most pertinent to our students, rather than focusing on improved test performance, we are truly being accountable as educators.

References