

CHAPTER 20



Legislating Competence

*High-Stakes Testing Policies and Their Relations
with Psychological Theories and Research*

RICHARD M. RYAN
KIRK W. BROWN

The development of competence in schools is an increasing focus of national concern in countries across the globe. This concern is fueled by the fact that educational outcomes, broadly considered, are linked with the health and economic well-being of nations. Beyond the obvious economic and health value of schooling to the individual person, the general expansion of education within a nation is associated with a host of outcomes, from reduced mortality and fertility to increased economic productivity and positive social change (Sen, 1999).

Because of the importance of the development of competence, governments are also increasingly attempting to legislate ways to enhance educational opportunities and outcomes. Yet much controversy exists about the appropriate ways governments can stimulate improved schools and greater academic achievement, and what kind of improvements in achievement are actually

meaningful for the health and economic well-being of a nation. This issue is international and occupies headlines from Great Britain to South Korea.

In the United States, state and federal government policies aimed at obtaining greater "accountability" and "higher standards" have especially stimulated controversy. These recent policy initiatives attempt to improve school performance through *high-stakes testing* (HST). Specifically, high-stakes policies represent a two-pronged approach to reform. The first prong entails increased testing to gauge how students, teachers, and schools are performing relative to each other, and relative to the *standards* that government agencies determine all students should meet. The second prong carries the motivational component: This testing has teeth. The attainment of standards is motivated or enforced by *high stakes* in the form of rewards and punishments, such as

financial incentives and job security for educators, and grade retention versus promotion for students. HST reform has become, in short order, the most dominant pressure in America's public schools and is rapidly reshaping teaching practice and curricular contents across the nation.

What is most interesting about this approach to reform, for the purposes of this volume, is that HST policies reflect particular theories of motivation and achievement. Specifically, high-stakes reform approaches represent a view of competence promotion and teaching that reflects an operant theory of motivation (Kellaghan, Madaus, & Raczek, 1996) and a view of educational outcomes that is more closely aligned with those espousing performance goals rather than mastery or learning goals (Deci & Ryan, 2002); that is, the governmental policy is founded on the idea that making rewards and punishments more salient and contingent on test score outcomes is the most appropriate and effective way of ensuring greater student effort and learning, and more effective teaching. As such, this social policy enacts a behavioristic motivational philosophy and represents a natural experiment in the social psychology of competence. It is a policy that suggests that high-quality educational motivation is a function of external incentives, a view that at least some psychologists support (e.g., Eisenberger, Pierce, & Cameron, 1999; Hidi, 2002).

In contrast, several theories in contemporary motivational psychology predict that attempting to enhance achievement in schools through such external controls will yield some highly negative results, based on the properties of the type of motivation it incites. In particular self-determination theory (Deci & Ryan, 1985; Ryan & Deci, 2000) explicitly predicts important costs of implementing such an approach to motivating competence in public schools. Similarly, some tenets of modern goal theories (e.g., Dweck, 1991; Nicholls, 1984; Elliot, 1999) also suggest potential costs of a focus on demonstrating performance outcomes. Thus, what is scientifically engaging about the social policy debate and implementation is that results of reform should be interpretable, in accord with the varying predictions of these psychological models. What is so-

cially engaging about the debate are the relative costs and benefits to children.

In this chapter, we examine HST reforms in the United States precisely because they illustrate the impact that social policy can have on institutional practice, and the relations (or absence of them) between empirically based research in psychology and education, and governmental policies. We highlight the nature of these test-driven reforms, the legislation surrounding them, and both the theoretically predicted impact and the current empirical data on their effectiveness and consequences. We then discuss the seeming divorce between political reforms and current empirical research in the psychology of competence and education.

To presage some conclusions, our review suggests that, to date, HST has not, in general, produced positive outcomes. Nonetheless, both the positive and negative data that have been obtained can be readily interpreted using the principles outlined in extant theories of motivation. In line with operant theory, and the general recognition of the power of contingent rewards to control behavior, high-stakes policies do indeed change behavior. They lead to increased district, school, and teacher activities intended to raise test scores. In fact, some of the behaviors that these contingencies incite are part of the problem, such as "teaching to the tests," elimination of developmentally enriching activities that are not likely to be tested, manipulation of targeted standards, and "push-outs" of potentially low performers from the pool of test takers. In line with self-determination theory (e.g., Deci & Ryan, 2002; Ryan & LaGuardia, 1999) and some perspectives on performance-focused motivation (e.g., Midgley, Kaplan, & Middleton, 2001), these high-stakes reforms are yielding a variety of collateral or unintended negative consequences, especially in areas involving persistence and quality of learning. Among the concerns is that HST is typically "one size fits all," requiring all students, regardless of their backgrounds, learning differences, and rates of development, to jump the same evaluative hurdles simultaneously. This approach potentially lowers the ability of schools to optimally challenge students of different talents and achievement levels, and it is of special concern regarding students with disabilities. An-

other concern is the problem of transfer: Rises in high-stakes test scores do not appear to generalize to other indices of improved achievement (e.g., other achievement measures). This poor generalizability is not necessarily due to the invalidity of the tests, but rather to the criterion contamination caused by their high-stakes implementation. The rewards and punishments that prompt an urgency to raise test scores lead to a narrowing of teaching, and therefore learning, and foster classroom dynamics that tend to decrease student motivation and engagement, as well as teacher morale and creativity. Perhaps more importantly, because HST neither provides a good basis for intrinsic motivation nor offers students optimal challenges (because the standards and methods of demonstrating performance are the same for all), reforms based on HST have been associated with increased school dropouts. These dropouts are especially salient among those already at risk, including the urban poor, students with special needs, and those for whom English is a second language—the very children whom many HST advocates have said they do not want to leave behind.

THE HIGH-STAKES TESTING MOVEMENT

There is little argument that gathering information and providing feedback about performance in educational settings is important for maintaining student and teacher motivation, and for informing educational policy (Linn, 2000; Shepard, 2000). Indeed, feedback regarding outcomes is recognized as a critical feature in improving the function of any organized system (Carver & Scheier, 1998). The function of assessment in gathering information, however, has additional impacts when the outcome data are linked with contingent rewards and punishments, as is the case in HST.

HST has been advocated as a means of motivating students and teachers alike to put in more effort, and thereby raise student achievement (Oakes, 1991; Finn, 1991). Policies instituting HST have taken on varied forms, but the common denominator in such initiatives is that state or federal governments mandate standardized testing of all students and then administer rewards or

sanctions based on the results. Students, teachers, and schools that improve or do well are rewarded, and those that decline or do badly are punished. For students, HST results can be the basis for promotion versus retention, and in some states, failure on a single indicator can result in the denial of a high school diploma. Teachers in schools that perform well may get cash bonuses, while those in other schools are reprimanded or derogated. For the schools, the comparative student performance average can result in increases versus cuts in school budgets, and in some cases, poor student performance may result in administration changes or even school takeovers by the state. When the stakes get high for administrators, local officials can even add to the stakes. For example, schools have offered cash prizes, parties, exemptions from finals, scholarships, candy, and awards to high-scoring students (Keller, 2000). School superintendents have been given personal cash bonuses when scores in their districts improve. However, the principal incentive at the administrative level is the public nature of high-stakes assessments. Schools and districts are publicly compared on their test scores, with the often explicit reasoning that pride or humiliation will be attached to the differences in score attainments. Accordingly, at all levels of educational systems, raising the stakes leads to increased attention to test scores because of the consequences attached to them.

A BRIEF HISTORY OF HIGH-STAKES TESTING

The modern HST movement has roots dating back to 19th-century England. Utilitarian philosophers such as Jeremy Bentham (1748–1832) and James Mill (1733–1836) formulated principles of motivation based upon hedonic principles and associationism that provided the foundations of what would become modern behaviorism (Rachlin, 1976). In applying these principles, they suggested the systematic use of rewards and punishments to establish good learning habits in schools. The English Parliament was perhaps the first government to put HST into practice, passing numerous laws intensifying examination structures to ensure liter-

acy, including the Revised Code of Regulations (1862), which advocated a "payment by results" scheme that linked the funds awarded to schools to students' performance on the exams. Whereas the Code promoted a wider national school system, it also prompted a rigid narrowing of curricula and an escalation of teacher-centered drill- and repetition-focused instruction. Although the Code was eventually repealed, the ideas of "streaming" or segregation of students according to ability level, evaluation by exams, and the resultant conservative methods instituted by the British system in the 19th century continued into the modern era.

In the United States, the modern instantiation of HST begins with the controversial publication of *A Nation at Risk* in 1983. This document, authored by the National Committee on Excellence in Education, declared that a rising tide of mediocrity was threatening the United States and its ability to compete in the world economy. (Parenthetically, one should note that despite relative stability in achievement standings since 1983, U.S. workers in 2001 were second in the world in global competitiveness according to the World Economic Forum [2002] report). Although one might assume that reform to alleviate "mediocrity" could take any number of directions, the U.S. government's approach under President Reagan was to step up demands for a core curriculum, more homework, more discipline, and more "accountability" (e.g., performance-based pay for teachers and increased testing), not more resources for schools, in part because lawmakers sought reforms that could be easily understood and rapidly implemented. Within several years following the report, virtually all states adopted more stringent graduation requirements, and many added mandatory homework requirements. School days lengthened and extracurricular amenities shrank. Standardized testing and curricula, matched to what those tests could measure, burgeoned.

Echoing the spirit of these reforms, William Bennett, a politician and popular moralist, proclaimed that "accountability is the linchpin, the keystone, the sine qua non of the reform movement" (Toch, 1991, p. 205). The demand for accountability led quickly to a focus on tests and pressure toward better outcomes on them. Policymakers in

nearly every state implemented policies to assess educational standards, and in many of these states, high-stakes consequences were attached to these outcomes, presumably as an incentive-punishment system to motivate change. High-performing schools were to be rewarded and underperformers penalized. Thus, the implementation of policy followed a behaviorist paradigm in which contingent rewards were applied to motivate (and control) teachers and students.

Although there were disappointing results from this early round of HST and many well-documented negative effects (see review by Toch, 1991), the late 1990s saw a new infusion of investment in HST policies. Politicians and business groups lobbied for still greater accountability in public schools, and states increasingly developed tests by which to rank and reward schools based on standardized test scores. Some states, such as Texas, aggressively pursued HST policies throughout the 1990s, and in so doing showed increased scores on the specific tests that were the targets of rewards and sanctions (Haney, 2000). By the first year of the new millennium, nearly all states were using HST in an attempt to foster school achievement. Nearly all states now publish school or district report cards on targeted tests, with the explicit purpose of motivating schools through public pressure or ridicule. Nearly half of all states also provide financial rewards to schools that improve on tests, and threats of administrative change or takeover for those that decline. Many states are directly paying school administrators bonus cash awards when schools under their watch improve on test scores.

Finally, states have been increasingly creating high stakes for students, as well as administrators. The most common high stake is that grade passage versus retention, and ultimately graduation, is contingent on passing a state-administered test. The high stakes of grade retention on the basis of a single examination have been applied as early as the fourth grade (e.g., in Florida). It is explicitly assumed by HST advocates that this type of contingency leads students to work harder in school (e.g., Cheney, 1991; Shanker, 1993), a point contested by critics (see Kelleghan et al., 1996). At this point in time, more than half of all states have made grad-

uation from high school contingent on a standardized test performance.

A National Initiative: No Child Left Behind

In 2001, President George W. Bush succeeded in passing, with bipartisan support, landmark legislation entitled No Child Left Behind (NCLB). A stated goal of NCLB is to raise levels of achievement and close the performance gap separating middle-class from poor and underperforming minority students. The plan called for even more testing and more salient stakes for schools and students alike. Specifically, NCLB mandates annual testing in grades 3–8 in math and reading. According to the legislation, scores from such tests are to be used to determine improving and declining achievement, such that penalties and rewards can be attached to them at the level of schools and children. Schools must make steady progress every year toward raising achievement levels on these exams in each of five racial and ethnic subgroups, as well as among low-income students and those with limited English skills or learning disabilities. Failure to demonstrate improvement for *any* of these subgroups for 2 consecutive years results in a school being labeled *low performing*. According to NCLB mandates, schools deemed low performing must facilitate the transfer of students to better schools or provide private tutors for students. Schools that continue to be low performing beyond 2 years can have their administrators and staff replaced. Federal funding is made contingent on compliance with these mandates.

NCLB has many critics. Given the expectable, year-to-year deviations that occur in standardized test results, schools may frequently be categorized as low performing for what amounts to statistical issues rather than reasons of educational quality. However, such logistical concerns are not the ones most pertinent to a critique of HST as a strategy of reform. As noted, HST represents a motivational policy. Yet a number of contemporary motivational theories suggest that a host of unintended negative consequences will stem from the pressure and rewards used to externally control teaching and learning. These include narrowing of curricula, teaching to the test, less creative

teaching, more superficial and nontransferable learning, more controlling behavior at all levels of power, more withdrawal of effort from at-risk students, and increased dropout rates. We turn first to these theoretical predictions, and then to a review of the accumulating empirical findings on the use of HST.

THEORETICAL PERSPECTIVES ON HIGH-STAKES TESTING

High-Stakes Testing as an Operant Approach

HST is based, at least implicitly, on a behaviorist view of student and teacher motivation. By putting contingent reinforcements on outcomes, the policy presumably increases efforts and behaviors associated with improvement; that is, HST advocates reason that whatever behaviors schools adopt to enhance test scores will be reinforced and selected for, whereas those associated with lower scores will be extinguished and, in the case of poor-performing schools, selected out. Not only will the behavior of teachers change, so will that of students. According to Shanker (1993), strong consequences attached to test scores will provide students with “the incentive to work hard and achieve because they know something important . . . is at stake” (p. 7).

The historical link between HST and behaviorism has deep roots. As previously noted, behaviorism emerged from a blend of British associationism and a hedonic view of human motivation, in which learned behaviors were always a function of external controls that punish or reward. It follows from this perspective that educators should utilize these external forces in regulating learning.

This approach to motivation was integral to the work of perhaps the most influential of all behaviorist educators, E. L. Thorndike. The central principle of Thorndike’s theory of learning, which he called *connectionism*, was his *law of effect*, which states that if a behavior is followed by a satisfying consequence, it is more likely to occur in the future under similar conditions. Conversely, if a behavior is followed by an unsatisfying consequence, its probability of recurrence will wane. A second principle was that of

frequency: The more frequently an association is repeated, the more likely it is to recur in similar conditions. Together, these "laws" of learning underwrote educational practices focused on the use of external reinforcements, coupled with practice, drill, and repetition. Although these techniques have characterized conservative approaches to education across history (see Ryan & Lynch, 2003), connectionism gave them a specific theoretical rationale.

Thorndike was also an advocate of testing. As he stated, "Testing the results of teaching and study is . . . the sine qua non of sure progress. It is the chief means to arousing . . . the instinct for achievement" (1962, pp. 65–66). However, interestingly, Thorndike was also cautious about how such tests should be used. As he states: "Great care should be taken in deciding anything about the fate of pupils, the value of methods, the achievement of school systems and the like from scores made in a test" (p. 156).

Thorndike's behaviorism was influential in education for several decades but eventually gave way to the "radical behaviorism" of B. F. Skinner. Skinner similarly advocated the systematic application by teachers of consequences, principally positive reinforcements, to induce learning. Skinner also promoted the idea of "programmed learning," which viewed instruction not as based in relationships or interests, but rather in a well-structured and systematic application of contingent reinforcements.

Today conservative educators continue to advocate the use of rewards to control learning, both at the classroom and school system levels. Behaviorists argue that teaching is most effective when based on control through reinforcements. For example, behaviorists Cameron and Pierce (1994), in the context of reporting a now discredited meta-analysis (see Deci, Koestner, & Ryan, 1999), argued that "teachers have no reason to resist implementing incentive systems in the classroom" (p. 397). At a political level, this theme is echoed loudly. Chester Finn has argued that "the problem is that academic success yields such few rewards [*sic*] and indolence brings few penalties" (1991, p. 120). He, and a broad array of conservative spokespersons, have argued that putting rewards and penalties behind the test scores will effectively alter the behavior of both

teachers and their students. This type of thinking has deeply influenced recent educational reforms in several nations focused on HST. In this view, instruction should be driven by measurement, and the outcomes of measurement should be the basis of rewards and sanctions for both teachers and learners (as discussed in Popham, 1983).

Our interpretation of the HST movement as reflecting an operant strategy has one very important caveat. Operant theory has always been focused on making rewards contingent on target *behaviors*. The twist in the HST movement is that its advocates apply contingent rewards and sanctions to *performance outcomes*; that is, rather than rewarding valued behaviors, such as student effort or work habits, contingencies are instead applied to test outcomes, the control over which is often questionable, especially for at-risk students. Similarly, rather than rewarding excellent teaching activities and approaches, schools are rewarded or sanctioned on their test score results. This practice is not in line with the fundamental tenets of the operant viewpoint. Indeed, we believe that the focus on performance outcomes, rather than on behaviors that students and teachers have direct control over, is one of the features of HST that lead to reinforcement of the wrong behaviors.

This focus on outcomes does find affinity from some theorists who focus on goals as motivating forces in behavior. Among those perspectives that could be aligned with HST-based reforms is the goal theory approach of Locke and Latham (1990), who argue for a high-performance model in which demanding goals are linked with both internal and external rewards to maximize organizational efficiency. Although they developed their model in application to industry, they suggest its generalizability to schooling, arguing that the high-performance model of difficult goals associated with rewards for success "should be made part of our schools as well as our work organizations" (p. 268). Advocating this linkage between measurable outcomes and performance-contingent reinforcements would seem to be fully congruent with the HST approach. A similar advocacy of applying contingent rewards to performance outcomes has also been forwarded by Hidi and Harackiewicz (2000), whose perspective on performance goals we

review in discussing theories of mastery and performance goals.

Organismic Perspectives on Learning

A very different view of what motivates learning and competence can be gleaned from what has sometimes been called the "liberal perspective," and sometimes the "organismic perspective," in which learning is seen as an inherent or intrinsic tendency of the person (Ryan & Lynch, 2003). In this tradition, the desire to learn is seen as a natural or basic tendency of humans. Learning is growth. However, like all growth, this inherent initiative or tendency requires support and nutriment. The result is a process (rather than outcome) focus, in which nurturance, mainly in the form of warm relationships, optimal challenges, and supports for autonomy and interest, are the most common elements.

Throughout history, educators embracing this liberal view have argued that students are not optimally motivated by external controls, but rather by support of their inherent tendency to learn. In ancient times, this view was espoused by Quintilian, who recognized that learners of different ages and types have distinct needs and interests, and held that curriculum and methods should be tailored accordingly. He deemphasized the then common use of punishment, instead stressing the importance of making learning interesting and attractive. In the Renaissance, similar views were echoed by Comenius, who focused on the strategic importance of warm student-teacher relationships and enhancing students' interest in learning. Subsequently, Enlightenment philosopher Rousseau laid the groundwork for much modern thinking in the liberal vein, emphasizing children's curiosity and natural inclination to learn under supportive conditions.

Rousseau influenced generations of subsequent educators. Outstanding among them was the Swiss educator, Pestalozzi, who viewed the aims of education not as "imposing on the child fixed doctrines and alien concepts but in helping him to develop his own constructive powers" (Silber, 1973, p. 274). His method of education entailed, first and foremost, an atmosphere of emotional security based in a warm and caring relationship between teacher and child. He

advocated that knowledge be gained, when possible, through direct experience rather than through mere words passed from teacher to child. He also downplayed the utility of punishment and fear of evaluation, suggesting that if provided a secure base, the child's nature would lead to discovery and growth. Pestalozzi's philosophy was widely disseminated during the 19th century in Europe and the United States, and became a major influence on a diverse family of practitioners, including Froebel in Germany, and Montessori in Italy.

Finally, in the 20th century, Dewey (1938) emphasized the importance of cultivating interest and inquiry in crafting an education, rather than arbitrarily imposed educational tasks and goals. He stood, in this respect, in stark contrast to his behaviorist contemporary, Thorndike. In the realm of psychology, Rogers (1969) developed an influential perspective on teaching, stemming from his *person-centered approach*. He advocated a classroom experience that grows out of the authentic inquiry of the student. Rogers felt that the external locus of evaluation represented by traditional examinations and normative grading stifled the significant learning that grew out of a student-centered, responsive teaching environment. It was Rogers who faced off with B. F. Skinner in the 1950s and 1960s, debating the value of external control versus self-actualization in the enterprise of learning.

In summary, a long tradition of philosophy and psychology has argued against externally controlling techniques as the *via regia* to student learning. Instead, this tradition focuses on nurturing the natural inclination to learn, the diversity of learning abilities and styles, and the importance of students' developing their powers of self-evaluation. Importantly, the last few decades have seen the emergence of several empirically focused motivation theories that supply some support for this perspective.

Self-determination theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2000) is one such empirically based organismic perspective that views humans as intrinsically motivated to learn and develop competencies. However, the theory is centrally concerned with the conditions that support versus thwart these intrinsic propensities. SDT is thus particularly interested in the impact of events

such as evaluations, praise, and contingent rewards and punishments on behavior and learning.

Specifically, SDT highlights the fact that students' motivation to learn can vary in its relative autonomy, from behaviors motivated by external rewards and punishments (controlled motivations) to those that are energized by interests and values (autonomous motivations). Both evidence and theory based on SDT suggest that, to the extent that one's motivation is based on more autonomous motives, such as intrinsic motivation or well-internalized values, the more quality of learning, persistence, and affective experience are enhanced (Grolnick & Ryan, 1987; Ryan & La Guardia, 1999; Ryan, Stiller, & Lynch, 1994). On the other hand, SDT research has found that motivation based on more controlled motives, such as rewards or punishments (external regulations), or self-esteem-based pressures (e.g., ego involvement) is associated with lower quality learning, lessened persistence, and more negative emotional experience.

Because HST policies are based on the idea that rewards, punishments, and self-esteem-based pressures are effective motivators of learning, the principles of SDT apply (Deci & Ryan, 2002; Ryan & La Guardia, 1999). In what follows, we summarize the theoretical basis for those hypotheses as they relate to teacher and student motivation, and review some of the evidence supporting the validity of these hypotheses.

According to SDT the specific effects of external events such as evaluations or feedback on human motivation depend on the psychological meaning, or *functional significance*, that the events have for the recipient (Deci & Ryan, 1980, 1985, 2000). The theory specifies that the functional significance of an external event, such as a test score, a tangible reward, or praise from a teacher, can be informational, controlling, or amotivating. Events have *informational significance* when they provide effectance-relevant feedback in a noncontrolling way; that is, when an event provides individuals with specific feedback that points the way to being more effective in meeting challenges or becoming more competent, and does so without pressuring or controlling the individuals, it tends to have a positive effect on self-motivation. Events have *controlling sig-*

nificance when they are experienced as pressure toward specified outcomes or as an attempt to control the activity and effort of the individual. According to SDT, when evaluations have controlling significance, they may produce temporary compliance, but they ultimately undermine self-motivation, investment, and commitment in the domain of activity being evaluated. Finally, events have *amotivating significance* when the feedback conveys incompetence to the individuals or supplies no inner or outer rationale for acting. Evaluations or reward structures based on overly challenging standards, or that are perceived to be beyond the reach of the individuals, are thus amotivating: They undermine all motivation and lead to withdrawal of effort. Teaching that does not tap into a student's interests, or that does not supply a basis for the experience of relevance or meaning, can also foster amotivation.

Both experimental and field studies have supported these predictions concerning the impact of events such as feedback and rewards on subsequent motivation. Extensive reviews are available elsewhere, but a few examples are worth detailing. In experiments with rewards, Ryan, Mims, and Koestner (1983) showed that reward structures delivered in an informational manner did not undermine intrinsic motivation, but rewards used to pressure people toward a specified outcome did. In another demonstration, Ryan (1982) showed that students who were pressured to perform by stressing that outcomes reflected ability (an ego-involving induction) were subsequently significantly less likely to engage with the target task than were students who were induced to focus on the task itself rather than task outcomes. In an experiment conducted within an elementary school, Grolnick and Ryan (1987) had students engage in a reading comprehension task under three conditions. In the first, students were told they would not be tested at all. In the second condition, they were told they would be tested, but only to determine what kinds of ideas were learned, so there were no consequences for failure or success. In a third condition, students were told they would be tested and graded, and that the grade would be delivered to their classroom teacher. This third condition represented a controlling use of

evaluations. Results showed that the controlling evaluation condition promoted not only short-term, rote memory but also produced a significantly lower level of conceptual learning and knowledge integration than the two noncontrolling conditions. Evidence from these and related studies (e.g., Benware & Deci, 1984) indicates that when tests, evaluations, and rewards are used in controlling ways, they have negative effects on students' interest, motivation, and higher level cognitive outcomes in school. Classrooms studies have added to these findings by showing that when teachers are oriented toward being controlling (e.g., using evaluations and rewards), students are less intrinsically motivated, less desirous of challenge in school, and also less confident in their skills (e.g., Deci, Schwartz, Sheinman, & Ryan, 1981; Ryan & Grolnick, 1986).

How Performance Standards Affect Teachers

The finding that when teachers use controlling strategies and performance pressure to motivate students, the students become less self-motivated and less engaged in school, raises an interesting issue. What factors lead teachers to be controlling? One answer is that they may become controlling when they themselves are pressured to get children to perform. An experiment performed by Deci, Spiegel, Ryan, Koestner, and Kauffman (1982) addressed this issue. Participants simulated teachers with the task of helping students learn a cognitive-perceptual task. The teachers all had the same set of problems to work with and were given the same preparation. However, before entering the teaching session, one group was explicitly told that it was their job to make sure their students performed "up to high standards," whereas another group received no such pressure. The sessions were recorded and rated for differences in teaching strategies. The results showed that the participants who were explicitly pressured to produce high student achievement were more controlling and less supportive of students' autonomy. Specifically, teachers in the performance standards condition engaged in more lecturing, criticizing, praising, and directing—all techniques that have been shown to have a negative impact on students' interest in learning

and their willingness to undertake greater academic challenges. Flink, Boggiano, and Barrett (1990) followed up on this reasoning by examining a newly introduced school-based curriculum for elementary students across several schools. They showed that, as predicted, teachers pressed toward higher standards were more likely to engage in controlling instructional behaviors. In line with SDT, the more they did so, the more their students actually performed more poorly on objective test-score outcomes. This is consistent with a wide body of literature linking evaluative pressure with poorer performance in schools (Kohn, 1996; Ryan & Stiller, 1991), as well as dropout rates (Hardre & Reeve, 2003).

From the SDT perspective, creating a test-driven evaluative focus not only leads teachers to be more controlling but also leads students to be more externally regulated and/or ego involved in their motivational orientation. According to SDT, ego involvement is potentiated whenever a person's esteem is linked with attainment of specific outcomes (deCharms, 1968; Plant & Ryan, 1985; Ryan, 1982). Accordingly, ego involvement can motivate effort, just as rewards can. However, like most performance-contingent rewards, ego involvement is a controlling form of extrinsic motivation, and it runs the risk of undermining internal motivations based in value or interest. Furthermore, unless one is ensured of success when applying effort, ego involvement can have deleterious immediate effects. The more ego involving a context, the more many students, particularly the less confident ones, withdraw effort in order to reduce the diagnosticity of tests and thus protect their self-esteem (Martin, Marsh, & Debus, 2001). Additionally, even for students who try to do well, such evaluation-based motivations tend to foster more superficial and less integrative learning processes, thus debilitating long-term knowledge retention and growth (Golan & Grolnick, 1990; Grolnick & Ryan, 1987).

Beyond this, the evidence suggests that focusing parents' concerns on performance outcomes will lead them, like teachers, to use pressuring motivational strategies that will backfire, leading to lower achievement over the long term (Ginsburg & Bronstein, 1993; Grolnick, 2003; Grolnick, Gurland, Decourcey, & Jacob, 2002; Grolnick &

