T·R·E

Undermining quality teaching and learning

A self-determination theory perspective on high-stakes testing

RICHARD M. RYAN AND NETTA WEINSTEIN
University of Rochester, New York, USA

ABSTRACT

Using tests to compare nations, states, school districts, schools, teachers, and students has increasingly become a basis for educational reform around the globe. Although tests can be informative, *high-stakes testing* (HST) is an approach to reform that applies rewards and sanctions contingent on test outcomes. Results of HST reforms indicate a plethora of unintended negative consequences, leading some to suggest that HST corrupts educational practices in schools. Although there are many accounts of these negative results, SDT supplies the only systematic theory of motivation that explains these effects. In what follows we describe the motivational principles underlying the undermining effects of HST on teachers and learners alike.

KEYWORDS *education reform, high-stakes testing, self-determination theory*

ONE SIMPLISTIC SOLUTION TO IMPROVING SCHOOLS is to use a carrot and stick approach: Apply rewards and sanctions contingently on standardized test score outcomes, and assume this will motivate administrators, teachers, and students to improve. These ideas represent the core of *high-stakes testing* (HST) policies. HST reform strategies are being dicussed and implemented around the globe, from Canada to South Korea, and are exemplified in the *Education Reform Act* in Great Britain and the *No Child Left Behind* (NCLB) legislation in the USA. Under such policies the results of examinations are

Theory and Research in Education

VOL 7(2) 224–233 ISSN 1477-8785 DOI: 10.1177/1477878509104327

used to determine student advancement, and to provide criteria for rewarding high-performing schools or sanctioning those whose test scores falter.

HST reforms represent a *motivational approach* because they not only put an emphasis on test scores; they also implement strategies to enhance these outcomes through contingent rewards or sanctions. Accordingly, this article will discuss the motivational implications of HST-based reforms from the theoretical position of self-determination theory (SDT; Ryan and Deci, 2000). SDT has long argued that using controlling external contingencies to change behaviors or enhance outcomes is typically ineffective over the long term, and yeilds many hidden costs. Although many commentators today are recognizing the damages associated with HST policies, few have a theoretical or empirical basis for understanding these effects. SDT supplies both of these.

SDT is an empirically based theory that is primarily concerned with promoting students' interest in learning, growth in competencies, and well-being. Within SDT people are viewed as having inherent and deeply evolved propensities to assimilate knowledge and develop new skills. Yet SDT argues that these natural propensities can be either supported or undermined by social contexts. School and classroom strategies, including the use of grades, evaluations, rewards and external pressures, are thus of particular interest within SDT as they impact our human potentials to learn and develop.

SDT distinguishes between *intrinsic motivation,* doing an activity for its inherent satisfactions, and *extrinsic motivation*, doing an activity for its instrumental value. Within SDT, extrinsic motives are further differentiated into those that are heteronomously regulated or *controlled* versus those that are more self-regulated or *autonomous* (see Niemiec and Ryan, this issue) SDT-based research has consistently demonstrated that more autonomous forms of motivation are associated with a host of positive outcomes from greater academic perfromance, creativity, and persistence, to enhanced learner wellness.

In terms of social contexts, SDT suggests that autonomous motives, and the energy and engagement associated with them, are supported by contexts that enhance experiences of *autonomy, competence,* and *relatedness*. In this view, the effects of classroom events such as examinations, teacher feedback, or the introduction of a new curriculum on students' motivation are determined by the *functional significance*, or meaning, of these events with respect to these three basic needs (Ryan and Deci, 2000) Similarly, the meaning of policies that reward or punish teachers or schools also have a functional significance – they will shape the type and focus of consequent administrator and teacher motivation. Specifically, the functional significance of any event can be either *informational, controlling,* or *amotivating*.

Events, including tests, can be experienced as *informational* when they provide noncontrolling feedback that the learner or teacher can utilize in

becoming more effective. According to SDT, to the extent that an event is experienced as informational, it tends to have a positive impact on self-motivation, enhancing both feelings of competence and autonomy. Events have a *controlling* functional significance when they are perceived as pressure toward a specified outcome, or as an attempt to control behavior (Deci et al., 1999). Tests, especially when connected with rewards and sanctions, can be experienced as controlling (Ryan and Brown, 2005). Although controlling events may prompt immediate compliance, people tend to exert the least effort required to gain rewards or avoid punishments, and a side effect is often diminished self-motivation, investment, value and the performance enhancements that stem from these. Controlling motivational strategies have been emprically shown to foster more superficial forms of learning, and to undermine more autonomous and engaged forms of motivation (Ryan and La Guardia, 1999). Finally, external events are experienced as *amotivating* when they convey incompetence or helplessness. Tests that are too challenging or result in highly negative feedback tend to discourage rather than inspire further effort.

The functional significance of HST has not been considered in most implementations. Yet linking performance outcomes with rewards or sanctions is one way to enhance their functional significance as controlling and/or amotivating. Strong empirical evidence suggests that such a linkage probably undermines both complex learning, as well as interest in students (Deci et al., 1999), and at a systemic level, to lower teacher morale and educational innovation (Ryan and Brown, 2005). Moreoever, because high-stakes assessments are intended to be uniform, or 'one size fits all', they are not optimally challenging for most individuals or school populations (Kohn, 2000a), and thus can readily have an amotivating functional significance.

Defenders of HST policies claim that these reward and sanction policies simply represent an effective use of reinforcements, in accord with behaviorist theories (Finn, 1991). Yet, as Ryan and Brown (2005) argued, classical operant theory (Skinner, 1953) applies reinforcement contingencies to targeted *behaviors*. HST practices, in contrast, apply contingent consequences to *outcomes* rather than behaviors. Ryan and Brown suggested that a danger with this outcome focus is that a wide variety of potential behaviors, both desirable (e.g. changes in instruction) and undesirable (e.g. teaching to the test, narrowing of curriculum, cheating) can be equally 'reinforced' insofar as they produce desired outcomes. Moreover, these policy makers seem unaware of the limits of operant thinking, which assumes an absence of inner motivation in learners and teachers. Because of this, operant methods too often undermine or fail to nurture these valued inner resources, such as interest and intrinsic motivation.

Substantial evidence supports SDT's predictions of how feedback and external evaluations can have different functional significance, and thus differing effects on motivation (Deci et al., 1999). Experiments have repeatedly demonstrated that rewards or feedback delivered in a controlling manner undermine intrinsic motivation and deeper forms of learning. For example, Grolnick and Ryan (1987) examined the controlling use of a test in an elementary school setting. Students were told to learn specific text passages because they would be tested and graded (controlling condition) or told that they would be tested, but only to identify what was learned (informational condition). It was found that the controlling use of the test resulted in less depth of processing and less conceptual integration. Students in the non-controlling, informational condition, in contrast, evidenced higher levels of conceptual learning and reported more interest and enjoyment for material. Similar results have been found in numerous studies (Ryan and Brown, 2004; Ryan and LaGuardia, 1999).

EGO-INVOLVEMENT

HST suporters such as Finn (1991) explicitly want to activate not only a desire to improve scores, but also a fear of failing. Sanctions are a salient force for students and teachers alike. It is often this threat of sanctions rather than rewards of success that are most notable to those subjected to HST, especially in high-poverty schools. Indeed, Miner (2000) notes that, in actual implementations of HST, punishments are enacted twice as often as rewards. Teachers, students, and administrators often experience such policies as primarily 'shame-based' motivators, focused on publicly comparing schools, and on threatening administrators and teachers (Nichols and Berliner, 2007). Within SDT, conditions in which one's reputation or self-worth are contingent on performance are referred to as *ego-involving*, which is considered a controlled form of motivation (Niemiec and Ryan, this issue). Ryan (1982) demonstrated that, when subjected to ego-involving climates, people report less interest, more pressure, and less desire to engage in an endeavor beyond what is needed to protect self-esteem. Numerous studies have since supported these hypotheses. Like other controlled forms of regulation, ego-involvement undermines intrinsic motivation, enhances anxiety, and leads to more impoverished learning. These findings highlight the fact that although controlling regulatory styles such as ego-involvement can lead to 'motivation', they also exact high collateral costs. HST potentiates ego-involvement for both educators and learners, and its negative effects.

EFFECTS ON TEACHERS

Just as HST can undermine student engagement, so too can it undermine best teaching practices. Placing controlling contingencies on teachers has been predicted within SDT to yield more controlling styles of teaching (see Pelletier et al., this issue). For example, Deci et al. (1982) designed a simulation study in which teachers were asked to instruct students. All teachers had the same preparation but, just before the teaching session, one group was told that they were to 'make sure the students performed up to standards'. A second group received no such instructions. Ratings of the teaching sesssions revealed that those explicitly instructed to produce high standards were more controlling: they engaged in more lecturing, criticizing, praising and directing – all techniques that have been shown to have a negative impact on students' interest and volition. Flink et al. (1990) similarly showed that teachers pressed toward higher standards in actual school settings were more likely to engage in controlling instructional behaviors. To the extent that they did so, their students performed poorly on objective test outcomes. Although paradoxical to the motivate-by-pressure crowd, these results are expected by SDT, and consistent with a wide body of literature linking evaluative pressure with poorer performance (Ryan and LaGuardia, 1999) and higher dropout (Hardre and Reeve, 2003). Consistent with SDT, such contexts also negatively impact teachers' experiences, leading to more job stress and burnout (Dworkin, 2001).

Predictably, HST has impacted drop-out rates (Orfield et al., 1999). Recall that SDT argues that nonoptimal challenges result in decreased or impoverished motivation and also lower persistence. Because HST policies apply a single standard to all, many learners are overchallenged, and others are bored. Moreover, because of the pressure posed by HST on administrators to improve school rankings, there is incentive for schools to rid themselves of students who could potentially drag down scores. Such strategies entail recategorizing low-achieving students into special programs, or encouraging them to pursue degrees elsewhere (such as by attaining General Educational Development tests: GEDs). In line with this, Haney (2000) found that exclusion rates explained score gains in Texas where HSTs were a strong focus. Other practices include preventing students from passing on to a grade where high-stakes milepost assessments are given, a practice linked to an increased risk of school drop-out (Clark et al., 2000). Some policies in the United States attempt to counter these practices by applying sanctions to schools, based not only on test outcomes, but also on attrition rates. The dual pressures have led districts to distort not only score reports, but also drop-out statistics, often

leading to great discrepancies between reported drop-outs and the actual 'disappearing rate' of students. Clark et al. (2000) verified this link between attrition rates and the use of HST, reporting that in Texas, where graduation is based on exit exams, black or Hispanic students were three times more likely to drop out, even controlling for background factors.

SDT also predicts that HST will have deleterious effects on the *content* of instruction. SDT argues that controlling rewards or contingencies typically lead people to become extrinically focused, and thus to take the shortest route to the specified end. If contingencies are focused on test outcomes, SDT suggests that HST will incite excessive test preparation activities, 'teaching to the test', and a narrowing of the curriculum to material expected to be on tests. HST thereby reinforces harmful behaviors such as culling low performers before testing, misreporting or distortion of test results, and controlling rather than supportive teaching climates to drill students toward higher scores on targeted tests. In accord with these expectations, studies have verified that undesirable behaviors are being reinforced under HST policies, from cheating at all levels of the school system (Kohn, 2000b; Nichols and Berliner, 2007) to more classrrom time spent 'drilling' test information to enhance scores (Jones et al., 2003).

HST policies foster the use of a more standardized curriculum within schools, and less pluralism in education. HST advocates claim that such uniformity insures that all students are receiving the same quality of education. Yet a standardized or 'one size fits all' assessment, and the curricula aligned with it, will lead some students to be underchallenged, some overchallenged, and few optimally challenged. For students with language or learning barriers, such standard assessments not only are inappropriate, they also demoralize. For the gifted they are irrelevant, limiting, and boring. HST therefore crowds out individualized and responsive education. Motivationally, this uniformity is clearly the result of an increased focus on test preparation rather than a concern with quality of education. For example, McNeil and Valenzuela (2000) found that teachers under HST reported realigning their instruction to focus on topics expected on the targeted exams. Significantly more time was spent on test-taking strategies rather than substantive issues. These trends were especially evident in schools serving less affluent students (Moon et al., 2003).

Yet another question has been whether increases on HST scores 'generalize' or transfer to nontargeted measures of acheivment. Klein et al. (2000) found that, in Texas, where test scores on high-stakes exams have increased, so too has the amount of test-focused instruction and preparation. Yet gains on targeted exams did not result in parallel improvements on the National Assessment of Education Progress (NEAP), a no-stakes indicator of learning. Amerein and Berliner (2002) collected test scores from 18 US states with

strong HST policies and found similarly that, compared with the nation as a whole, HST policies did not lead to improved performance on the SAT, ACT, or NEAP. Nichols et al. (2006) reported similar findings across 25 states. Thus, there is little evidence that the pressure created by HST policies reliably leads to genuine gains in learning.

In sum, SDT highlights aspects of HST that can have multiple deleterious effects. The HS in HST policies represent a controlling intervention that tends to undermine autonomous motivation and encourage a more narrowed and impoverished approach to teaching, and a diminished focus on best practices.

SUMMARY AND CONCLUSIONS

HST represents a motivational strategy that, because it is controlling and extrinsic in character, often raises targeted test scores in the short term while producing a plethora of unintended negative long- term consequences. Nichols and Berliner (2007) discuss these issues in terms of *Cambell's law*: the idea that attaching serious consequences to any indicator increases the probability that its meaning and utility will be corrupted. While that *names* the problem, it does not explain how and why such corruption occurs. Teaching to the test, narrowing of curricula, crowding out of enriching student activities, test preparation resulting in poor generalization of gains, and the other corruptions we described, are *motivated* phenomena – they occur because of the controlling nature of HST policies. These effects of HST can all be predicted from SDT, and indeed have been for over two decades.

From an SDT perspective it is not tests, per se, that are the problem, but rather the stakes contingently attached to them. Assessments can have informational value, especially when used along with other performance indicators. Tests can be useful in documenting schools needing more resources, in comparing curricula, and in identifying populations requiring more intensive or alternative approaches. For students, they can help identify gaps in fundamental knowledge, or lack of progress in specific competencies. However, when high stakes are attached to tests, their informational value becomes corrupted. HST policies do 're-form' educational practices by placing excessive emphasis on outcomes, and a corresponding inattention to the optimal processes and best practice methods of educating our young.

From our view, schools are not factories with an aim of producing a standardized product, but rather contexts to foster human development (Ryan and Lynch, 2003). Like all developmental processes, progress must be nurtured rather than force-fed, and that requires an understanding of the nutriments through which true growth occurs. In this brief article we have addressed

only the problems with HST approaches to reform, rather than the positive appproaches that an SDT framework would advocate. Many of these are discussed in other articles within this special issue. Instead of attempting to threaten or seduce schools to improve through external contingencies, an SDT approach would work with stake holders, including parents, administrators, teachers, and students, to identify barriers to change and the goals to which they aspire, and to actively empower and support change from within. Not only does this result in greater engagment and knowledge, it also models the democratic processes and responsibilities we think schools should prepare all students to assume.

REFERENCES

Amrein, A.L. and Berliner, D.C. (2002) 'High-stakes testing, uncertainty, and student learning', *Education Policy Analysis Archives* 10(18). Available at: http://epaa.asu.edu/epaa/v10n18/ (accessed January 2008).

Clark, M., Haney, W. and Madaus, G. (2000) 'High-stakes testing and high school completion', *National Board on Education and Testing Policy* 1 (3). Available at: http://www.bc.edu/research/nbetpp/publications/v1n3.

Deci, E.L., Koestner, R. and Ryan, R.M. (1999) 'A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation', *Psychological Bulletin* 125: 627–68.

Deci, E.L., Spiegel, N.H., Ryan, R.M., Koestner, R. and Kauffman, M. (1982) 'The effects of performance standards on teaching styles: The behavior of controlling teachers', *Journal of Educational Psychology* 74: 852–9.

Dworkin, A.G. (2001) 'Perspectives on teacher burnout and school reform', *International Education Journal* 2(2): 69–78.

Finn, C. (1991) *We Must Take Charge: Our Schools and Our Future*. New York: Free Press.

Flink, C., Boggiano, A.K. and Barrett, M. (1990) 'Controlling teaching strategies: Undermining children's self-determination and performance', *Journal of Personality and Social Psychology* 59: 916–24.

Ganesh, A. and Berlinger, D. (2006) 'What do positive high correlations of teacher's ranking with students' actual standardized test scores mean?'. Paper presented at the annual meetings of the American Educational Research Association, San Francisco.

Grolnick, W.S. and Ryan, R.M. (1987) 'Autonomy in children's learning: An experimental and individual difference investigation', *Journal of Personality and Social Psychology* 52: 890–8.

Haney, W. (2000) 'The myth of the Texas miracle in education', *Education Policy Analysis Archives* 8(41). Available at: http://epaa.asu.edu/epaa/v8n41

Hardre, P. and Reeve, J. (2003) 'A motivational model of rural students' intentions to persist in versus drop out of, high school', *Journal of Educational Psychology* 95: 347–56.

Jones, M.G., Jones, B.D. and Hargrove, T.Y. (2003) *The Unintended Consequences of High-stakes Testing*. Lanham, MD: Rowman and Littlefield.

Klein, S.P., Hamilton, L.S., McCaffrey, D. and Stecher, B. (2000) 'What do test scores in Texas tell us?', *Education Policy Analysis Archives* 8(49). Available at: http://epaa.asu.edu/epaa/v8n49/

Kohn, A. (2000a) 'Burnt at the high-stakes', *Journal of Teacher Education* 51: 315–27.

Kohn, A. (2000b) *The Case against Standardized Testing: Raising the Scores, Ruining the Schools*. Portsmouth, NH: Heinemann.

McNeil, L. and Valenzuela, A. (2000) *The Harmful Impact of the TAAS System of Testing in Texas: Beneath the Accountability Rhetoric*. Cambridge, MA: Harvard University Civil Rights Project.

Miner, B. (2000) 'Testing: Full speed ahead', in K. Swope and B. Miner (eds), *Failing our Kids: Why the Testing Craze Won't Fix our Schools*, pp. 114–17. Milwaukee, WI: Rethinking Schools, Ltd.

Moon, T.R., Callahan, C.M. and Tomlinson, C., (2003) 'Effects of state testing programs on elementary schools with high concentrations of student poverty – Good news or bad news?', *Current Issues in Education* 6(8). Available at: http://cie.ed.asu.edu/volume6/number8/

Nichols, S.L. and Berliner, D.C. (2007) *Collateral Damage: How High-Stakes Testing Corrupts America's Schools*. Cambridge, MA: Harvard Education Press.

Nichols, S.L., Glass, G.V. and Berliner, D.C. (2006) 'High-stakes testing and student achievement: Does accountability pressure increase student learning?', *Education Policy Analysis Archives* 14(1). Available at: http://epaa.asu.edu/epaa/v14n1/.

Orfield, G., Losen, D., Wald, J. and Swanson, C. (1999) 'Losing our future', in J.P. Heubert and R.M. Hauser (eds), *High-Stakes Testing for Tracking, Promotion, and Graduation*. Washington DC: National Academy Press.

Ryan, R.M. (1982) 'Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory', *Journal of Personality and Social Psychology* 43: 450–61.

Ryan, R.M. and Brown, K.W. (2005) 'Legislating competence: The motivational impact of high-stakes testing as an educational reform', in C. Dweck and A. Elliot (eds), *Handbook of Competence*. New York: Guilford Press.

Ryan, R.M. and Deci, E.L. (2000) 'Intrinsic and extrinsic motivation: Classic definitions and new directions', *Contemporary Educational Psychology* 25: 54–67.

Ryan, R.M. and La Guardia, J.G. (1999) 'Achievement motivation within a pressured society: Intrinsic and extrinsic motivations to learn and the politics of school reform', in T. Urdan (ed.), *Advances in Motivation and Achievement*, Vol 11., pp. 45–85. Stanford, CT: JAI Press.

Ryan, R.M. and Lynch, M. (2003) 'Motivation and classroom management', in R. Curren (ed.), *A Companion to the Philosophy of Education*, pp. 260–71. Malden, MA: Blackwell Publishing.

Skinner, B.F. (1953) *Science and Human Behavior*. New York: Macmillan.

BIOGRAPHICAL NOTES

RICHARD M. RYAN is a Professor of Psychology, Psychiatry, and Education at the University of Rochester. He is a Fellow of the American Psychological

Association, an award-winning teacher and researcher, earning a Phi Delta Kappa award twice for research contributions in education. Ryan is also Editor-in-Chief of the psychological journal *Motivation and Emotion*. He is a widely published researcher and theorist in the areas of human motivation, development, and psychological well-being, having published over 300 articles and chapters. He is the co-developer of self-determination theory, a widely researched general theory of human motivation that has been applied to such areas as education, health care, organization, behavior, psychopathology, and sport and exercise. Correspondence to: Richard M. Ryan, Clinical and Social Sciences in Psychology, University of Rochester, R. C. Box 270266, Rochester, New York 14627, USA. [email: ryan@psych.rochester.edu]

NETTA WEINSTEIN is a doctoral student in clinical psychology at the University of Rochester and is part of the human motivation program. Her research examines supports for, and consequences of, autonomous functioning, focusing primarily on outcomes relating to emotion and interpersonal regulation.

Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing

Richard M. Ryan and Netta Weinstein

Theory and Research in Education 2009; 7; 224
DOI: 10.1177/1477878509104327