TEST-BASED TEACHER EVALUATIONS: ACCOUNTABILITY VS. RESPONSIBILITY

Chloé Bolyard
Miami University

Gert Biesta contends that managerial accountability, which focuses on efficiency and competition, dominates the current political arena in education.[1] Such accountability has influenced states' developments of test-based teacher evaluations in an attempt to quantify teachers' efficacy on student learning. With numerous state policies requiring the use of student test data to influence decisions regarding teacher pay, tenure, and contract renewals, the significance of this discussion should not be overlooked.[2]

During the 2013–2014 school year, Ohio implemented the Ohio Teacher Evaluation System (OTES), linking half of a teacher's evaluation to student growth. This paper first briefly summarizes OTES and discusses its implied notions of accountability and responsibility. The second part of the paper critiques OTES as a form of disciplinary power with various consequences and potential responses for educators. Therefore, the purpose of this paper is to examine the notions of accountability and responsibility as related to current teacher evaluation policies in an attempt to stimulate a conversation about two questions: What does a discourse of high-stakes accountability within OTES imply about a teacher's responsibility in education? What are the consequences and subsequent implications for educators?

Ohio's Teacher Evaluation Policy (OTES)

Robert Wagner asserts that accountability efforts have attempted to bridge public interests and the performance of schools with a general agreement on the following: assessing the quality of schools must occur by carefully examining the school's output (i.e., student learning); learning can be measured as a form of cost-effectiveness; the stakeholders have a "right" to know about the school's costs and outputs; and accountability will provide a stimulus for enhanced performance.[3] Larry Cuban extends Wagner's

---

[1] Gert Biesta, *Good Education in an Age of Measurement: Ethics, Politics, Democracy* (Boulder, CO: Paradigm, 2010), 2.

[2] Bruce Baker, Joseph Oluwole, and Preston Green, "The Legal Consequences of Mandating High Stakes Decisions Based on Low Quality Information: Teacher Evaluation in the Race-to-the-Top Era," *Education Policy Analysis Archives* 21, no. 5 (2013): 3–4.

[3] Robert Wagner, *Accountability in Education: A Philosophical Inquiry* (New York: Routledge, 1989), 1–2.

accountability discussion by noting the historical shift in school accountability. Cuban notes that school boards have been accountable to voters since the inception of tax-supported public schools. Identifying 1965 as the turning point, Cuban defines the pre-1965 accountability movement as one focused on schools efficiently providing students access to adequate materials, staff, and buildings. Since 1965, however, accountability efforts have defined good schools "as ones that have efficiently used their resources to yield improved students' academic achievement as measured by test scores."[4] Ohio adopted one such policy that targeted teacher performance.

In 2009, House Bill 1 directed Ohio's Educator Standards Board to create evaluation systems for teachers and principals. From 2009-2011, various stakeholders collaborated to draft the Ohio Teacher Evaluation System (OTES). Using other states' evaluations as exemplars, OTES was designed to provide ongoing assessment and feedback to support improved practice.[5]

Under OTES, fifty percent of a teacher's evaluation is comprised of his/her performance on standards, while the other half consists of student growth measures. Using the *Ohio Standards for the Teaching Profession*, performance is rated on a rubric through a variety of means (i.e., conferences, walk-throughs, observations, and professional growth or improvement plans). Based on a teacher's performance, s/he will be rated as accomplished, skilled, developing, or ineffective.

Student growth is defined as "the change in student achievement for an individual student between two or more points in time."[6] Student growth measures are figured for teachers depending on the subject and grade taught using Ohio's state assessments or Local Education Agency (LEA) approved assessments (i.e., Student Learning Objectives, shared attribution measures, and vendor assessments). Based on student growth, a teacher will be rated as falling below expected growth, meeting expected growth, or exceeding expected growth.[7] If a teacher is rated ineffective based on performance and/or student growth, then s/he is placed on an individual improvement plan designed by her/his principal. After a reassessment of the teacher's performance, the teacher may resume the regular evaluation cycle if desirable improvement was achieved. If the teacher's performance does not meet expected performance levels, then further improvement plans may be implemented or steps may be taken to recommend dismissal.[8] Furthermore, local boards of education will

---

[4] Larry Cuban, "Looking Through the Rearview Mirror at School Accountability," in *Holding Accountability Accountable: What Ought to Matter in Public Education*, ed. Kenneth A. Sirotnik (New York: Teachers College Press, 2004), 27–28.

[5] "OTES Model Packet," *Ohio Department of Education*, July 2013, http://education .ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation -System/Teacher-Performance-Ratings.

[6] Ibid., 5.

[7] Ibid.

[8] Ibid.

determine how evaluations impact retention and promotion decisions and removal of ineffective teachers.[9]

Notions of Accountability

To better understand the possible consequences and implications of OTES, a brief discussion on some notions of accountability is necessary. According to Wagner, accountability means to give an explanation for something. One might infer that being accountable simply stated means being answerable to. The degree of accountability thus varies from explaining to justifying one's actions.[10] For Wagner, two types of accountability exist: moral and legal accountability. When one is morally accountable, s/he exercises certain obligations that require ethical considerations based on moral principles. Legal obligations, in contrast, use law to enforce certain actions by removing the freedom to choose otherwise.[11] Fundamental to Wagner's discussion of moral and legal accountability is his assertion that laws often result from a lack of moral sensitivity, reiterated by Zygmunt Bauman.[12] Similar to Wagner, Cuban defines educational accountability as "fixing responsibility—either moral or legal or both—and providing relevant information on the efficiency and effectiveness of schools to those who make informed decisions."[13]

For Bruce Charlton, Wagner's definition of accountability as "being answerable to" sums up the first of two notions of accountability.[14] The first relies upon mutual responsibility. However, Charlton contends that a second form of accountability is associated with a technical-managerial meaning, with its roots in fiscal monitoring (e.g., auditing financial expenditures).[15]

Biesta posits that managerial accountability in education has resulted from a changing ideological framework, specifically the rise of neoliberalism and capitalism. These changes have resulted in a "*reconfiguration of the relationship between the state and its citizens*" from a *political* relationship to an *economic* relationship.[16] This changing relationship has caused schools to focus on quality assurance, raising standards, and outcomes. Instead of encouraging mutual responsibility according to Charlton's general definition of accountability, OTES positions the relationship of the state and the school above the relationship between the school and the community. With two vastly

---

[9] "State Board of Education Approved Teacher Framework," *Ohio Department of Education*, September 2013, http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System.

[10] Wagner, *Accountability in Education*, 7–8.

[11] Ibid., 79–95.

[12] Biesta, *Good Education,* 61.

[13] Cuban, "Looking Through the Rearview Mirror," 21–22.

[14] Biesta, *Good Education,* 51–52.

[15] Ibid.

[16] Ibid., 53.

different notions of accountability, the rhetoric in education has caused individuals to associate accountability with responsibility, which positions those arguing against accountability as proponents of irresponsible action.[17]

Responsibility

One may assume that responsibility is a necessary condition of accountability, but being responsible does not always necessitate an obligation to account. This causes one to ask: What is responsibility? Wagner draws a distinction between causal responsibility and expectational responsibility.[18] One's performance or nonperformance is responsible for something or someone in the former. For example, a parent is causally responsible for his/her actions, which affect his/her child's welfare. Expectational responsibility is associated with roles that agents assume. Responsibility to one's profession, community, or family derives from the expectations associated with each role.[19] In education, teachers are professionally responsible for the children under their care.[20] A teacher, therefore, feels that his/her actions impact the actions of his/her students.

By evaluating a teacher through his/her students' growth, OTES implies that teachers are responsible for test scores. To hold teachers responsible for test scores assumes that they have control over the circumstances that affect a student's performance on a test. As Wagner notes, "Not only is this assumption superficial and out of touch with reality, it presents a major obstacle to the improvement of student performance and to the achievement of many outcomes that are important in education."[21] I argue that this assumption oversimplifies a complex phenomenon. Wagner illustrates the variety of factors associated with student performance, specifically citing the 1966 Coleman report.[22] Conducted by James Coleman, the large-scale empirical study found that socioeconomic status, home life, and peer culture had a greater impact on student learning than did curriculum and instruction.[23]

If the underlying purpose of accountability policies is to improve student performance, how can this be achieved by holding only one party responsible?[24] This is not to say that teachers do not affect student performance, but to be held accountable implies that they are exclusively

---

[17] Ibid.

[18] Wagner, *Accountability in Education*, 49–53.

[19] Ibid.

[20] Michael Gunzenhauser, *The Active/Ethical Professional: A Framework for Responsible Educators* (London: Continuum, 2012), 66.

[21] Wagner, *Accountability in Education*, 127.

[22] Ibid.

[23] William Schubert, et al., *Curriculum Books: The First Hundred Years*. 2nd ed. (New York: Peter Lang, 2002), 195.

[24] Wagner, *Accountability in Education*, 127.

responsible. Assuming that each factor is associated with student performance, teachers should not be held exclusively responsible for growth when legislators and taxpayers are not held responsible for the resources they provide.

By holding teachers accountable for student performance, OTES assumes that teachers are the ones to blame when students perform poorly. This becomes problematic based on the aforementioned description of the various factors associated with student performance. Since many parties share a role in contributing to a student's success, then it would seem logical that shared responsibility to improve student success is necessary. The language of blame, however, decreases the likelihood for collective action.[25] As Young contends, the language of blame impedes collaborative discussions about desired outcomes.[26] Blame produces resentment that results in defensiveness. Individuals focus on identifying those responsible and thus lose sight of the change they want to achieve.[27] In education circles, it is common to hear blame being passed around: "If only parents would be more involved in their child's learning," or "If teachers weren't so lazy," or "If unions didn't interfere with the dismissal of ineffective teachers," or "If taxpayers would just pass a levy." In basing teachers' evaluations, in part, on student test scores, OTES reinforces the notion of blame by implying that teachers are the only ones responsible for student test scores. Such blame language prevents the conversation from moving toward one that acknowledges a sense of shared responsibility.

OTES not only relies on teachers' responsibility for student growth, but its measurements illuminate only those teacher responsibilities that its creators value. Biesta calls for a normative validity by questioning "whether we are indeed measuring what we value, or whether we are just measuring what we can easily measure and thus end up valuing what we (can) measure."[28] While the scope of this paper is not to discuss OTES's implied aims of education, it is nevertheless important to acknowledge that which is not measured in relation to teacher responsibility. Consider, for instance, that one of the aims of education not measured is to promote a democratic society.[29] What would a teacher's responsibility be to reach this aim? Harry Brighouse

---

[25] Iris Marion Young, *Responsibility for Justice* (New York: Oxford University Press, 2011), 114.

[26] Ibid. Although Young uses blame when discussing structural injustices (i.e., poverty, unequal housing), I feel that an extension of her contribution is appropriate for this discussion.

[27] Ibid.

[28] Biesta, *Good Education,* 13.

[29] John Dewey, *Democracy and Education: An Introduction to the Philosophy of Education,* (New York, NY: The Free Press, 1916), 87. According to Dewey, "A democracy is more than a form of government; it is primarily a mode of associated living, of conjoint communicated experience." Dewey characterizes a democratically constituted society as one that recognizes mutual interests and allows for the development of new habits through free interaction between social groups.

contends that this is done, in part, by teaching what it means to be a well-functioning citizen.[30] Good citizenship is characterized as having three central components: to be law-abiding unless laws are broken in pursuit of justice; to participate actively in politics; and to politically participate with an attitude of respect and openness to others' views.[31] Although such aims of education may be identified as valuable, OTES certainly reveals what it values in education based on what it measures, and likewise what it does not value based on what it does not measure.

ACCOUNTABILITY OR RESPONSIBILITY

"Accountable for" and "responsible for" are conventionally parallel. Numerous responsibilities hold individuals accountable. For example, parents are responsible for the care of their children, and they are held accountable for neglect.[32] However, responsibility often exists without accountability. Parents are not held accountable for refusing to help their child attend college. In education, teachers are responsible for students but held accountable for student test scores. Nel Noddings distinguishes between accountability and responsibility by arguing that accountability is directed upward in a hierarchy (i.e., one is accountable to a higher authority); while responsibility "points downward in the hierarchy" (i.e., one is responsible for those whom one serves in a position of authority).[33] Michael Gunzenhauser echoes this distinction by asserting that the difference between accountability and responsibility depends on which relation is of primary importance. For teachers with professional responsibility, the relationship with the student is primary. In accountability policy, however, the primary relation is between the state and the school. By placing the primary relation between these entities, significant personal relationships become secondary.[34]

This shifting relationship becomes problematic, as schools are not directly accountable to stakeholders. Instead, the school is accountable to the state. Biesta contends that the public supports accountability so that schools will be answerable to the public. Accountability policies instead remove members of the public from the accountability loop. Parents, then, become consumers of their children's educations and are unable to participate in any public, democratic discourse about education.[35]

---

[30] Harry Brighouse, *On Education* (New York: Routledge, 2006), 62–73.

[31] Ibid.

[32] Wagner, *Accountability in Education*, 55.

[33] Nel Noddings, *When School Reform Goes Wrong* (New York: Teachers College Press, 2007), 39.

[34] Gunzenhauser, *Active/Ethical Professional*, 67.

[35] Biesta, *Good Education,* 52–56.

OTES is a high-stakes accountability policy in that a teacher's evaluation and subsequent improvement plans determine the future of his/her employment. Consequently, student growth, making up half of a teacher's evaluation, becomes critically important. High-stakes policies serve as a form of power to force individuals to comply by causing teachers to fear the consequences of inadequate performance and being labeled as low performers.[36] It becomes difficult to ignore the effect this has on the day-to-day decisions teachers make. Teachers would strive to be "normal."

Gunzenhauser draws on Michel Foucault's theory of disciplinary power to describe how accountability policies serve to self-discipline through normalization and surveillance.[37] "Discipline occurs through normalization, through which a norm is named, reinforced, and refined."[38] In accountability policies, assessments use a norm against which to evaluate students. Therefore, students are no longer treated as subjects but as objects;[39] they become objects of normalization. Tests are the tool for normalization by categorizing students, teachers, and schools. Assessments create conditions that have turned pedagogy into a science; teachers' professional philosophy of education is supplanted by enacting a scripted curriculum. Teaching becomes routinized.[40] This seems to be the paradox of OTES, a tool intended to improve teacher practice that causes regression to the mean, as educators must enact the desired norm through implementation and adherence to a scripted curriculum.

It may seem that OTES has only semi-normalized student performance by evaluating progress rather than proficiency. Progress is preferable as it considers the individual starting points of each student. However, OTES has essentialized how students grow academically by determining what counts as adequate growth. This norm becomes problematic, as growth is difficult to achieve for students at the high and low ends of the spectrum. Teachers are incentivized to focus on the students in the middle who are more likely to meet their expected gains.[41]

Accountability policies also act as a form of disciplinary power through the use of surveillance.[42] In Ohio, evaluations are required to be entered into an online database, the electronic Teacher and Principal Evaluation

---

[36] Gunzenhauser, *Active/Ethical Professional*, 82–98.

[37] Ibid.

[38] Ibid., 83.

[39] Dewey, *Democracy and Education*, 29. Objects, according to Dewey, are given meaning according to their use. Therefore, in the case of viewing students as objects under test-based accountability policies, teachers act toward students in a way that is perceived to promote the achievement of a certain test score.

[40] Gunzenhauser, *Active/Ethical Professional*, 82–98.

[41] Diane Ravitch, *The Death and Life of the Great American School System: How Testing and Choice are Undermining Education* (New York: Basic Books, 2010), 270.

[42] Gunzenhauser, *Active/Ethical Professional*, 82–98.

System (eTPES). The stated purpose of eTPES is to calculate student growth measures, calculate the summative rating, and generate a record for the state of Ohio and the U.S. Department of Education.[43] eTPES is an example of how accountability policies exercise control through surveillance.

Foucault creates a metaphor with the panopticon, a method of monitoring inmates from a centralized guard station which causes inmates to discipline themselves by knowing a guard could be watching them at any moment.[44] By considering Foucault's panopticon, one can see how surveillance via data entry (i.e., eTPES) would cause teachers to discipline themselves to produce student growth consistent with the norm (i.e., expected gains). Self-discipline becomes problematic when it removes the freedom to choose between alternative possibilities of action. Again, teaching becomes routinized and pedagogy becomes regimented: What teaching practices will produce the scores needed to meet the norm?

Unintended Outcomes of Accountability Policies

The demands of high-stakes accountability policies in various states have resulted in numerous unintended outcomes. These outcomes include narrowed curriculum, cheating scandals, and teacher burnout. In addition, accountability polices have caused a culture of fear and anxiety to replace a culture of collaboration.[45]

Moreover, accountability policies have become a philosophical problem. According to Gunzenhauser, a philosophy of education "addresses *why* we educate so that we make better choices about *who*, *what*, *where*, *when*, and *how* we educate."[46] Community conversations should determine the purpose and aims of education. OTES, while developed by a group of stakeholders, applies a statewide evaluation system instead of locally created evaluations. Such accountability policies have narrowed education by using assessments as instrumental ends that overshadow students as ends-in-themselves.[47] A default philosophy of education that views test scores as dominant supplants teachers' professional philosophies of education, which allow for creativity and innovation. Such a default philosophy "results from a lack of reflective, engaged dialogue by educators and school communities about their goals and practice."[48] The normalizing pressures of assessment

---

[43] "Frequently Asked Questions: Electronic Teacher and Principal Evaluation System (eTPES)," *Ohio Department of Education*, February 2014, http://education.ohio.gov/ Topics/Teaching/Educator-Evaluation-System/District-Educator-Evaluation-Systems/ eTPES-Help.

[44] Gunzenhauser, *Active/Ethical Professional*, 82–98.

[45] Ravitch, *Death and Life*, 269.

[46] Gunzenhauser, *Active/Ethical Professional*, 32.

[47] Ibid., 34–35.

[48] Ibid., 9.

cause teachers to forgo their personal philosophies of education in order to comply with the default philosophy.

Implications for Ohio Educators

As OTES is enforced, where does this leave teachers who run the risk of being normalized? Teachers have one of three options: comply with the normalizing test, resist, or subvert. To comply means that teachers accept a default philosophy of education over their own professional philosophies. This allows the student-teacher relation to become secondary to the relation between the school and the state. The test becomes the subject of education, as students are viewed first as a test score to be maximized and second as a child whose learning is to be nurtured. Teachers may feel a professional responsibility to prepare students to have a successful future and may believe a child's test score will open or close various doors of opportunity. To resist would be seen as acting irresponsibly by inhibiting a student's access to opportunities (e.g., grade promotion). To resist may even result in lower test scores, which would label her/him as an ineffective teacher and may result in the loss of her/his job. To subvert, a teacher tries to align her/his actions with her/his own philosophy while still achieving the desired norm of student growth within the accountability system.

What would it look like if despite all of the foregoing considerations teachers chose to resist normalization? To resist would mean that a teacher protects the student-teacher relationship as primary. In so doing, the teacher treats the student as the subject of education, valuing students' and their learning as ends-in-themselves rather than as means to a test score. By resisting the normalizing effects of test-based accountability, teachers may act in a way that is congruent with their professional philosophies of education. Such an educational philosophy eschews the narrowness of scripted curricula, and, instead, embraces a broadened curriculum that fosters alternate perspectives and possibilities. Resistance also requires perspective; teachers resist by viewing the test score as one of many measures of a child's progress. Just as teachers are subjected to the normalizing technologies of high-stakes accountability policies, they are also in a position to exert normalizing power over their students by viewing all students in the same way.[49] Such a position requires teachers as ethical professionals to continually reflect on the power that they exercise over the students in their care.

Conclusion

The purpose of OTES to support improved educator practice leading to enhanced student performance is a noble endeavor. In addition, proponents of this policy might contend that OTES was created democratically by allowing multiple stakeholders' input. These proponents might further assert that the student growth measure of OTES is a drastic improvement over alternative

---

[49] Ibid., 82–98.

models that require students to meet a proficiency norm. Each of these claims is a strength of OTES when compared to the alternative, but I argue that each of these appraisals is overshadowed by the foregoing critiques of OTES.

My purpose with this paper has not been to argue that evaluation policies should be eliminated but rather to stimulate a discussion about the implied responsibilities and potential ethical implications for educators under the requirements of OTES. Certainly, evaluations have the potential to provide teachers with feedback to improve their practices. However, as OTES only measures student and thus teacher performance in content areas, is this measure consistent with what stakeholders hold to be the aims of education? Moreover, if half of a teacher's evaluation is based on student performance, over which a teacher exercises only partial control, then where does the notion of shared responsibility come in? If this same aspect of a teacher's evaluation results in numerous consequences incongruent with a teacher's professional philosophy of education, then how can Ohio move toward an evaluation system that honors teachers as autonomous professionals?