
Responding to the Demands of Assessment and Evaluation in Catholic Education

Leanne M. Kallemeyn

Loyola University Chicago

Public schools are experiencing a new era of assessment and evaluation with the implementation of state accountability systems and No Child Left Behind. How can Catholic schools respond by recognizing the legitimacy of evaluation and assessment, while also critically examining its appropriateness? To help address this question, this article contains a brief history of how Catholic schools have accommodated evaluation and assessment demands, followed by a critical examination of standards-based accountability systems and high-stakes testing in relation to the identity and current milieu of Catholic schools. It also presents an expanded understanding of assessment and evaluation, which draws heavily on qualitative research paradigms.

There is perhaps not a Catholic elementary school principal or teacher who has not at times wondered about the effectiveness of the educational program being offered in her school. Superintendents, too, have often reflected on the need for some method of evaluating their elementary schools, and have sought some means of ascertaining just how successful Catholic education is. (Catholic University of America Department of Education, 1949, p. 1)

The questions of effectiveness, success, impact, outcome, or influence of Catholic education are persistent questions that researchers and Catholic educators have addressed for decades (Convey, 1992, 2002). Educational evaluation and assessment¹ have emerged as critical practices to help address questions of effectiveness. The purpose of this paper is twofold. First, it is to understand better the current demands for evaluation and assessment in Catholic education. Historically, Catholic schools have formally integrated assessment and evaluation practices into their schools, paralleling

1 Evaluation is “the systematic assessment of the worth or merit of an object” (Joint Committee on Standards for Educational Evaluation, 1994, p. 3). Merit refers to the quality of the object itself. Worth refers to its value to society or the extent to which it fulfills a great need. In this paper, the object is Catholic schools. Assessment is a “process of gathering, describing, and quantifying information about performance,” typically a student’s performance (Center for Research on Evaluation, Standards, and Student Testing, 1999, Glossary). While evaluation and assessment are distinct practices, they share much in common. The arguments in this paper are applicable to both evaluation and assessment.

developments in public schools. Public schools are experiencing a new era of assessment and evaluation with the implementation of state accountability systems and *No Child Left Behind* (NCLB). Most notably, legislation attaches consequences to poor performance on assessment, such as closing a school. For this reason, assessment and evaluation results have a heightened status or authority in judging a school's effectiveness. How will Catholic schools respond to such changes in assessment and evaluation activities? Will Catholic schools also heighten the authority of assessment and evaluation? To what extent and in what ways are these assessment and evaluation activities consistent with the core values of Catholic schools?

Second, the purpose of this paper is to articulate a response to the current demands for evaluation and assessment in Catholic schools that recognizes the legitimacy of evaluation and assessment, while also critically examining its appropriateness. In order to do so, I will first discuss how Catholic schools have historically accommodated evaluation and assessment demands. I will then critically examine the current practices of standards-based accountability systems and high-stakes testing in relation to the identity and current milieu of Catholic schools. In the process, my intent is not to criticize and advocate against the use of assessment and evaluation practices. These practices have been, will be, and ought to be adapted for Catholic schools. Rather, my intent is to raise issues that may facilitate critical accommodation; more specifically, acceptance of thoughtful standards-based reform efforts and rejection of high-stakes testing practices. It is also to expand an understanding of assessment and evaluation, given the core values and identity of Catholic schools, which draws heavily on qualitative research paradigms and methodologies.

Significant Demands for Assessment and Evaluation in Catholic Schools

Demands Emanating from the Context of Public Schools

The understandings and practices of evaluation and assessment in Catholic schools have developed alongside the dominant culture of assessment and evaluation in public schools. A major characteristic of current evaluation and assessment practices in public schools is a heightened authority of student assessment results for decision making. Student assessment scores are used to guide classroom instruction and student grouping, to provide accountability to standards or benchmarks, and, in some cases, are used to make high-stakes decisions, such as a teacher's salary or whether to close a school. Currently, state level accountability systems and NCLB legislation heavily influence the evaluation and assessment practices in public schools.

These reforms are rooted in the evidence-based practice movement, which began in medicine in the 1990s and has grown substantially since then, including into the field of education (Thomas & Pring, 2004; Trinder, 2000). Evidence-based practice refers to the production and use of objective evidence to inform the decision making of the public, policy makers, parents, teachers, school administrators, and so on. One impetus for the emergence of evidence-based practice has been a gap between research and practice, and thus concern that teachers' and administrators' practices are ineffective. Evidence is intended to provide a stimulus and the necessary information to reform classroom-level and school-level instructional practices. Another impetus relates to funding. Unlike increases in federal funding for social programs in the 1960s during the Great Society Movement, recent concerns with deficit spending mean that policy makers are often looking for ways to cut spending. Borrowing from the business sector, one way of doing so is to have standards, indicators, or benchmarks of effectiveness that provide evidence to determine where funding cuts may be justified. When this evidence is made public, the intent is also to provide accountability to the taxpayer and consumer for the product.

The incorporation of evidence-based practice in education has its roots in the standards movement in the 1990s (Haertel & Herman, 2005; Linn, 2006). At the federal level, the *Improving American's School Act of 1994* and the Clinton Administration's *Goals 2000: Educate America Act* stimulated the growth in standards-based reform and accountability systems. By the 1999–2000 school year, 48 states used a state assessment as the primary indicator of school performance, and the other two states required the use of a locally chosen assessment (Goertz & Duffy, 2001). In all 50 states, results of these assessments were publicly reported. In 33 states, policies outlined consequences for schools based on students' performance on assessments, such as school improvement planning, monitoring or other intervention strategies, and monetary and recognition rewards. Linn (2000) argues that this current wave of assessment reform has three features. First, it emphasizes the development and use of ambitious content standards that inform assessment and accountability practices. Second, it simultaneously sets demanding standards and requires that all students are included in assessments. Finally, the assessments have high-stakes implications for schools, teachers, and sometimes students.

Alterations and expansions of these existing state accountability systems occurred as the result of NCLB legislation of 2001 (Linn, Baker, & Betebenner, 2002). The law required states to have content standards in reading and mathematics with tests linked to the content standards in Grades 3

through 8, and required that at least 95% of students participate in the tests. Most notably, the law required that states set adequate yearly progress (AYP) targets, so that by 2014 all students will be meeting standards. These targets need to be met within subgroups of students, in addition to the overall population. Schools that fail to meet their targets in 2 consecutive years are identified for improvement.

Currently, state accountability and assessment policies and NCLB do not directly affect Catholic education. Private schools are not expected and, in some circumstances, not even allowed to participate in state assessment systems. They are also not expected to meet regulations under NCLB. To the best of my understanding, no state report cards for Catholic schools are generated and published based on standards and assessments; thus, Catholic schools do not experience consequences from the state or federal level for their performance.

This being said, these policies still may have a significant impact on Catholic schools. For example, in 2008, the Iowa state legislature approved a new state curriculum mandate that requires state-accredited, non-public schools to adhere to the Iowa Core Curriculum (Robelen, 2008). Whether this legislation is found to be constitutional and upheld is yet to be seen, but this example does illustrate some impetus for Catholic schools to adopt state standards, a necessary step prior to assessments. Also, Catholic schools may receive funding through the *Individuals with Disabilities Education Act* (IDEA) or NCLB in order to serve their students. The laws require public school districts to give students in private schools equitable opportunities to receive available resources. In a sample of Catholic schools, 80% reported having funding support for at least one student in at least one of the 11 programs in NCLB, the highest rate of participation among various types of private schools (Christensen et al., 2007).

While the previous example may be limited to funding for individual students, charter school and voucher policies introduce government funding for private schools on a larger scale. In 2006, the Florida Supreme Court ruled that the state of Florida's Opportunity Scholarship program, a state-funded voucher program for private schools, was unconstitutional. According to an article in the *National Catholic Reporter*, "a major argument against the program is a lack of accountability and standardization of testing, teacher credentials, and curriculum" (Gross, 2006, p. 8). One implication of such a ruling is that future iterations of voucher programs and charter school policies may require some sort of accountability system for participating Catholic schools. Given the support among the Catholic community for such policies,

pressures may exist to develop accountability systems for Catholic schools or participate in state accountability systems.

While direct influences of these policies are limited at this time, the greatest influence may come from the discourse surrounding assessment policies in public schools, which penetrates to Catholic educators. In order to be perceived as exhibiting good practices, Catholic educators may adopt practices from public schools. For example, an archdiocese may want every school to complete the National Catholic Educational Association (NCEA) Assessment of Catechesis/Religious Education (ACRE), or another faith formation assessment, in order to know which schools in the diocese are effective and ineffective in religious education. In other words, Catholic schools and dioceses may choose to adapt and develop their own evaluation and assessment practices in a manner similar to public schools.

Demands Emanating from the Context of Catholic Schools

In the past few decades, similar to the perception of failings within public schools, Catholic schools have their own perceptions of failure. Most notably, Catholic school enrollments have declined since the 1960s (Convey, 1992) and there is a lack of monetary and personnel resources to support Catholic schools (Harris, 2000), which has resulted in school closures—particularly in urban areas (Nelson, 2000). These circumstances heighten the importance of illustrating the value of a Catholic education to parishioners, the general public, and policy makers looking for alternatives to public education. In the same way that accountability through assessments is perceived as important for public education, accountability through assessments may also be perceived as important in Catholic education reform. For example, the United States Conference of Catholic Bishops (USCCB) issued *Renewing Our Commitment to Catholic Elementary and Secondary Schools in the Third Millennium* in 2005. In response to this statement, the Notre Dame Task Force on Catholic Education (2006) issued the report *Making God Loved, Known, and Served*, which outlines 12 recommendations for the University of Notre Dame to enact a renewed commitment to Catholic education. One of the 12 recommendations is to “build a national initiative for the academic improvement of Catholic Education” (p. 7). The Task Force further explains:

While many factors have contributed to enrollment declines in Catholic schools, we have learned that the lack, or perceived lack, of academic excellence in some Catholic schools has played a significant role in these declines...This initiative will invest deeply in the research, development, and implementation of

effective assessment, curriculum, and instruction in Catholic schools in the following ways:

- Conduct research on current best practices in curriculum and instruction.
- Provide professional development workshops for teachers, principals, and superintendents on curriculum development, instruction, and assessment.
- Provide a rubric for the collection of student outcomes data.
- Develop benchmark goals tied to data for student results.
- Communicate an assessment process for annually and publicly evaluating student outcomes along with protocols for the use of this assessment to improve curriculum and instruction.
- Invest in the expansion and effective use of technology in Catholic elementary and secondary schools.
- Create the Alliance for Catholic Education Press to facilitate the rapid and inexpensive publication of articles, books, and resource materials nationally. (p. 7-8)

One perceived need for reform in Catholic education is significant changes to instructional practices. The actions to enhance academic excellence listed above borrow heavily from the standards-based reform movements in public education as a means of facilitating these changes in instructional practices, particularly a rubric for the collection of student outcomes data, benchmark goals tied to student results, and an assessment process for annually and publicly evaluating student outcomes.

Another recommendation in the report is to “access public funds and resources for Catholic schools and their students,” which includes “school-choice programs, tuition tax credits and deductions, publicly-funded transportation and textbooks, loaned computers and technology, special-education resources, and so on” (Notre Dame Task Force, 2006, p. 12). Accessing public funds means interfacing with the dominant culture of educational policy that views accountability as essential for ensuring quality education, particularly through assessments. For this reason, Catholic schools may experience similar pressures for evidence. A position to support school-choice legislation may enhance the demands for evidence of Catholic school effectiveness. While past studies have suggested that Catholic schools are more effective than public schools, particularly among students in low-income communities (Convey, 1992), current research on this topic is likely to be important in these efforts.

A Historical Perspective: Assessment and Evaluation Practices in Catholic Schools that Accommodated Public School Practices²

Evaluation and assessment were an integral part of educational reform throughout the twentieth-century. In this section, I will highlight three ways in which Catholic schools actively developed and implemented evaluation and assessment: program evaluation during the Great Society Movement, self-evaluations for school improvement, and student assessments.³ In each of these areas, Catholic schools adopted processes or practices for doing evaluation and assessment similar to those of public schools. The major accommodation was the addition of evaluation and assessment for religious education. In other words, Catholic schools augmented *what* was evaluated and assessed in order to reflect the Catholic identity. Educational researchers and theorists have argued that assessment and evaluation activities define and represent who we are and what we are doing. Lee Cronbach and his colleagues (1980) claim “whatever the evaluator decides to measure tends to become a primary goal of program operators” (p. 5). Fortunately, history has demonstrated that evaluators of Catholic schools have chosen to do the challenging work of aligning their core values with evaluation and assessment practices. They have attempted to measure their primary goals. Catholic school supporters developed these assessments to represent what it means to be a Catholic school and in what ways students benefit from being in a Catholic school.

A second observation is that over the course of history, evaluation and assessment activities in Catholic schools have primarily been completed for the purpose of learning how to improve and further develop Catholic schools. Two fundamental purposes for evaluation and assessment are facilitating improvement and providing accountability (Alkin & Christie, 2004; Mark, Henry, & Julnes, 2000).⁴ Linn (2001, 2004) argues that one of the key issues in the development and use of assessment systems is the purpose of the system, and that the design of assessments is consistent with the purposes. As the Catholic education community continues to use, and possibly expand, assessment and evaluation to enhance academic excellence and acquire additional

2 This section is intended to be illustrative of historical practices of assessment and evaluation in Catholic schools rather than a comprehensive description.

3 Efforts have also been made to formalize teacher evaluations, or personnel evaluations, which I will not address in this paper.

4 Accountability refers to the “process of ‘giving an account’ or being answerable” (Alkin & Christie, 2004, p.14). This process raises two questions: To *what* is a school answerable (i.e., what services, indicators of quality, benefits to society matter?), and to *whom* is a school answerable (i.e., students, parents, community members, funders, policy makers, the public, staff members)? Accountability can have consequences or high stakes associated with it. Program improvement or development refers to efforts to provide timely feedback in order to modify, improve, and/or enhance a program. Such evaluation typically focuses on program processes rather than outcomes (Mark et al., 2000).

public funding, will facilitating improvement be the focus, or will there be an increased emphasis on accountability?

Program Evaluation During the Great Society Movement

The growth of program evaluation in the United States has been linked to the Great Society Movement when numerous social programs were implemented and expanded as part of the War on Poverty in the 1960s (Haertel & Herman, 2005; Shadish, Cook, & Leviton, 1991). For educational programs, such as the *Elementary and Secondary Education Act* (ESEA), student assessments were critical measures of program effectiveness. With increased funding to alleviate social ills, politicians and the public began to ask questions. Is this a good investment of funds? What are the best approaches to alleviating poverty? At the time, models for evaluation were heavily drawn from disciplines in the social sciences that aimed to use the scientific method of the natural sciences to understand the social world.

While Catholic schools did not necessarily benefit from federal funds (McAndrews, 2003), the first large-scale study of Catholic schools, *Catholic Schools in Action*, occurred from 1962–1966 (Neuwien, 1966). This study described the enrollment and staffing for a large sample of Catholic schools in the United States. It also undertook the first large-scale attempt to assess students' religious understanding and attitudes through the Inventory of Catholic School Outcomes instrument. These early efforts in program evaluation persisted. Convey (1992) provided a synthesis of the research and evaluation efforts from 1965 to 1992. Such efforts, both in Catholic schools as well as in public schools, were consistent with Donald Campbell's (1981) vision of an "experimenting society." In an experimenting society, people use objective, scientific evidence to identify and understand problems and then generate and implement solutions. These solutions, or reforms, are tested for their effectiveness through experiments. If a particular reform does not demonstrate an impact on solving the program, then it is terminated and a new solution is sought.

Self-Evaluations for School Improvement

Extensive efforts have been made to encourage self-evaluations of Catholic schools, which are similar to a long history of school improvement planning in public schools. Evaluation guides have been developed for both high schools and elementary schools in a similar manner. For illustrative purposes, I will focus on elementary school evaluations. A similar story could be told with high school evaluations.

In 1949, the Department of Education at The Catholic University of America published the *Criteria for the Evaluation of Catholic Elementary Schools*. This document presents standards or criteria for evaluating Catholic schools, which are organized in the following areas: philosophy and objectives of Catholic elementary school education, school plant (i.e., building, grounds, and facilities), administration and supervision, curriculum and courses of study, materials of instruction, and teaching and learning activities. For each of these areas, checklists and scoring instruments were developed. The criteria were “designed primarily to assist superintendents, supervisors, principals, and teachers in assessing the educational endeavors of their schools. It is hoped that these criteria will also be provocative of specific improvements in Catholic education” (p. 1). In other words, the aim of the *Criteria* was to “enlighten” (Weiss, 2004), fostering the gradual percolation of ideas from evaluation into a school by questioning taken-for-granted assumptions, evoking new ideas, and altering priorities. Cronbach and his colleagues (1980) have also used the term “educative,” meaning that school personnel and the community learn something new about what it means to be an effective Catholic school.

As a result of a large demand for *Criteria*, The Catholic University, through the NCEA, revised the criteria in 1965 (National Catholic Educational Association Department of Elementary Schools, 1965). Each area in the original document was more comprehensively developed than the previous version, particularly the area of curriculum. This document also reiterated that “[these] criteria [are] not intended as an instrument of inspection. Its primary purpose is self-evaluation with an aim toward improvement and enrichment of the school program” (p. 3).

At a time when public schools were responding to *A Nation at Risk* and pressures to reform public education, Catholic schools were doing the same. In 1983, the NCEA compiled a list of 23 school self-assessment instruments utilized in Catholic schools (Reck & Coreil, 1984). Subsequently in 1984, the NCEA published *Verifying the Vision: A Self-Evaluation Instrument for the Catholic Elementary School* (Reck & Coreil, 1984). It included the areas of philosophy, Catholic school as community, Catholic school in the community, teaching/learning program, and organizational services. In comparison to prior documents, this one provided greater attention to the process of doing self-assessments through three steps of awareness, assessment, and analysis. These three steps illustrated how evaluation and assessment practices have developed based on modernist assumptions and an “ideology of social scientific progress,” which claims that scientific rationality or reason is the best

form of knowledge, and that such knowledge can be used to direct societal improvement (Schwandt, 2002). Evaluation and assessment involves identifying a problem, developing a solution, and then evaluating the effectiveness of the solution. What we experience or observe in a systematic manner provides factual evidence that we can use to make rational decisions.

As recently as 1999, an updated version was published for high schools: *Validating the Vision: An Assessment Protocol for Mission Effectiveness, Institutional Accreditation, and Strategic Planning in Catholic High Schools* (Taymans, 1999). This assessment was organized around school community, climate and culture, curriculum and instruction, learning media resources, schedule, student services, student activities and athletics, professional development, governance and accountability, school finance, institutional advancement, and facilities, and school and community relations. This tool integrated evaluation and assessment with strategic planning.

These self-evaluations for school improvement encourage school practitioners and local community members to be involved actively in the process of doing evaluations in a collaborative manner. A central idea in evaluation and assessment theory is that the process of doing an assessment or evaluation can be as meaningful and influential, if not more so, than the results of the evaluation or assessment. How assessment and evaluation happens matters. Patton (1997) coined the term “process use,” and described how evaluators ought to plan activities in order to facilitate benefits for stakeholders through the process of doing the evaluation. For example, when teachers participate in developing an assessment of religious education, they need to articulate systematically what religious education entails and, thus, gain a better understanding of what it is. Collaboration, dialogue, and learning among teachers would not occur if an external researcher developed the assessment instrument. Also, one of the main findings in research on evaluation is that the more people are involved and invested in the process, the more likely they are to utilize and make changes based on evaluation findings (Cousins & Earl, 1992).

Student Assessments

Extensive effort has also been made to assess students. At the turn of the 20th century, E. L. Thorndike developed standardized tests in reading, language arts, arithmetic, spelling, and drawing, which began large-scale efforts to assess students (Haertel & Herman, 2005). Such assessments have been used to track or group students into various educational opportunities (i.e., norm-referenced tests); have students demonstrate minimum competencies (i.e., high

school exit exams); admit students to college (i.e., ACT, SAT); provide students with college credits (i.e., Advanced Placement [AP]); monitor student learning (i.e., student portfolios, classroom-based assessments); diagnose learning difficulties; and most recently, as the centerpiece of school accountability systems (i.e., National Assessment of Educational Progress [NAEP], state assessments).

In public schools, student assessments focus on the academic content areas, such as reading and math. Currently, Catholic schools also utilize assessments for academic content areas. While no systematic research has been conducted on the use of student assessments in Catholic schools, common assessments encountered in Catholic schools include Preliminary SAT (PSAT), SAT, PLAN (i.e., Pre-ACT), ACT, AP exams, Iowa Test of Basic Skills (ITBS), TerraNova, student portfolios, and various curriculum-based and teacher-developed assessments to monitor student learning and diagnose learning difficulties. While Catholic schools typically do not participate in state assessment systems, a sample of Catholic schools do participate in the NAEP.

More notably, due to the core philosophy of faith formation, Catholic school supporters have extensively developed assessments for faith formation. These efforts began in research studies in the 1960s, such as *Catholic Schools in Action* (Neuwien, 1966), and later developed into individual student assessments that can be routinely administered to students. In 1976, the first instruments—Religious Education Outcomes Inventory (REOI) and the Religious Education Inventory of Knowledge, Attitudes, and Practices (REKAP)—were implemented. In 1978, they were combined into the Assessment of Catholic Religious Education (ACRE; Convey & Thompson, 1999; Thompson & Philibert, 1982). Similarly, during the 1970s, there was also a rise of minimum competency testing in public schools for academic content areas. Currently, the NCEA ACRE has three levels that can be administered at Grades 11 or 12 (Level 3), Grades 8 or 9 (Level 2), and Grades 5 or 6 (Level 1). The intent of the NCEA ACRE is to assess the strength of religion programs by measuring “students’ religious knowledge, beliefs, perceptions and practices” (Convey & Thompson, 1999, p. vi) based on the Catholic Church’s catechetical directives. The religious knowledge domains include God: Father, Son and Holy Spirit; Church: one, holy, catholic and apostolic; liturgy and sacraments; revelation, Scripture, and faith; life in Christ: personal morality and Catholic social teaching; Church history; prayer; and Catholic faith literacy. The religious beliefs, attitudes, practices, and perceptions include relationship with Jesus, images of God, Catholic identity, morality, students’ concerns, relationship with others, and perceptions about school,

parish, and religious program. Today, approximately 2,500 Catholic schools utilize the NCEA ACRE to inform the development and improvement of their religious education programs. Some Catholic schools and archdioceses have also developed and implemented other assessments of student faith formation. Unlike public school student assessments, the NCEA ACRE and other similar assessments have not been formally integrated into accountability systems in order to make judgments about the quality of Catholic schools.

What Is an Appropriate Response to the Significant Demands in Assessment and Evaluation in Catholic Schools Today?

Standards-Based Reform Efforts Are Consistent with the Nature of Catholic Schools

Catholic schools have a reputation for having high standards for all students, and emphasizing the inclusion of all students. For example, a finding in Bryk, Lee, and Holland's (1993) research on Catholic high schools is that the core curricula and teachers in Catholic high schools maintain high expectations for all students in the school. Catholic high schools also have a strong value for serving the underserved. These values represent two of the three characteristics in standards-based reform movements, as described by Linn (2000)—ambitious content standards and requirements for all students to participate in standards-based assessments. A reason for standards and corresponding assessments with consequences in public schools is an effort to raise standards that teachers and schools are expected to meet. If these expectations already exist in Catholic schools, then is it necessary to impose them through assessments in Catholic schools?

Even if these expectations already exist in Catholic schools, they may benefit from reminders. The question becomes: What is the appropriate way to do this in Catholic schools? Is it necessary to do so through assessments? Or given the character and heritage of Catholic schools, is it possible to do so on the basis of Catholic identity? Will reminders of the philosophy, values, and beliefs about Catholic education serve these ends? How can we reinvigorate teachers and administrators to this mission? The USCCB's (2005) *Renewing Our Commitment to Catholic Elementary and Secondary Schools in the Third Millennium* exemplifies such an approach.

While high expectations that are characteristic of standards-based reform movements may be integral to the identity of Catholic schools, these standards may not necessarily be as explicit as they are in public schools. Evidence from standards-based reforms in public schools has consistently shown that when standards are linked to assessments, teachers and principals

adapt teaching and curriculum to what is covered on the assessments (Herman, 2007); in other words, they do provide an impetus for reforming instructional practices. Formally articulating what the high standards are for all students in Catholic schools may also be beneficial for communicating with external audiences, such as parishioners, dioceses, public school officials, policy makers, the general public, and so on. Ways of formalizing these standards may be through accreditation processes, adoption of state standards, or individual school or diocesan practices. As discussed previously in this paper, a characteristic throughout the history of assessment and evaluation practices in Catholic schools has included assessments of religious education. Since assessment does provide a stimulus for what educators address in classrooms and schools, then inclusion of religious education is essential for maintaining the identity of Catholic schools, even though such information may not be necessary in the charter school and voucher debates, or among other external audiences.

Articulating and agreeing upon what the standards are and what we should evaluate and assess is challenging. Efforts within Catholic education exemplify such challenges. That being said, they also exemplify that in particular times and places, agreement can be reached. On the one hand, there is an agreed-upon sense of what the aims of Catholic education are, for example, the USCCB's *Protocol for Assessing the Conformity of Catechetical Materials with the Catechism of the Catholic Church* (1998) and *Doctrinal Elements of a Curriculum Framework for the Development of Catechetical Materials for Young People of High School* (2008). They have been repeatedly articulated in the Catholic magisterium. The NCEA ACRE represents an agreed-upon assumption of what faith formation entails, based on the USCCB's (1998) *Protocol for Assessing the Conformity of Catechetical Materials with the Catechism of the Catholic Church*. On the other hand, Catholic educators inevitably emphasize different aspects of Catholic education. For example, McLaughlin, O'Keefe, and O'Keefe (1996) argue that conflicting interpretations of the identity of Catholic schools exist through two extracts of Church documents, the first from the General Congregation of Jesuits and the second from an American archbishop seeking to give direction to the schools in his local archdiocese. The authors conclude: "Both of these statements were written and promulgated at the same time in history on the same topic: the fundamental identity of Catholic educational institutions. The statements are not necessarily contradictory; nonetheless, diversity of texture and tone are unmistakable. This [is] reflected in the wide ranging debate about how the distinctiveness of the Catholic school is to be understood"

(p. 17). In the process of articulating standards, preserving the “diversity of texture” will be critical.

While Catholic schools may benefit from additional efforts to articulate learning standards, some cautions are critical as schools proceed. First, in a review of research on standards-based accountability systems, Herman (2007) concludes, “It is clear from the research that accountability is changing what gets taught, but whether the change represents real improvement in students’ opportunity to learn is moot” (p. 19). The narrowing of the curriculum due to testing seems to minimize a focus on complex, higher-level thinking. Second, standards-based reform efforts have a tendency to oversimplify the teaching and learning process. Evaluators have argued that educational accountability systems typically represent a narrow set of decontextualized outcomes and are not meant to be responsive to dynamic schooling contexts (Greene, 1999; Ryan, 2002). Finally, Abma and Noordegraaf (2003) argue that certain public sectors may benefit from performance measurement systems, or accountability systems, more than others. Performance measurement may be less suited for management when services depend on interactions with participants and are not routine tasks, which is characteristic of school administrators. Managers deal with a high level of ambiguity. “The managerial world is a disorderly world: Managerial working days are hectic and chaotic; issues are unstructured and complex; attempts to influence others are difficult; and the effects of managerial interventions are hard to predict. Most managers are aware of this” (p. 288). As a result, when assessment systems do not assist with decision making, a strong tendency exists for these practices to become bureaucratic rituals and routines that become void of the meaning that they initially set out to represent.

High-Stakes Assessment and Evaluation Practices Are Not Consistent With the Nature of Catholic Schools

As argued previously, Catholic schools have a strong history and current practices in using evaluation and assessment for school improvement, such as through school self-assessments. Such evaluation activities encourage the participation of teachers and administrators in the process and, thus, enhance the learning and impact of such activities. In this way, assessment and evaluation become a “process by which society learns about itself” (Cronbach et al., 1980). This learning includes better understanding of social problems, identification of potential solutions and future program directions, identification of what circumstances result in the most effective and beneficial program processes, building infrastructure and capacity, and so on. This purpose is important due to the contexts within which social and educational programs

are situated, which are complex, multifaceted, and dynamic (Patton, 1994). High-stakes assessment practices no longer emphasize improvement, development, and learning; rather, the fundamental purpose is accountability or monitoring. Using assessment for these purposes is not consistent with the identity of Catholic schools.

First, high-stakes assessment introduces a different means of developing trust between society and school professionals. In the context of high-stakes testing, society can “trust” that school personnel are appropriately teaching students if the test scores validate that students are learning. Such a notion of trust is different than the trust that develops through interpersonal relationships between parents and teachers that provide opportunities to discourse about the best interests of a particular student. Such a notion of trust is different than society recognizing that teachers are professionals; therefore, it is important to trust their professional judgment as to the most appropriate ways to teach students. High-stakes testing in education is an example of what Powers (1997) refers to as the “audit society.” He argues that “in the audit society, institutionalized trust, which differs from the trust of ordinary individuals, is bestowed on the auditor and is displaced from other organizational locations” (p. 173). This institutionalized trust erodes the trust of ordinary individuals. In *Catholic Schools and the Common Good*, Bryk, Lee, and Holland (1993) described how Catholic schools feel like “home” and are “communal.” Such a schooling context has a strong sense of trust among ordinary individuals. Introducing high-stakes testing in Catholic schools may erode this central characteristic of Catholic schools.

Second, Catholic educators also recognize the moral dilemmas inherent in underenrolled, underfunded, understaffed, and underachieving schools. High-stakes assessment systems oversimplify these dilemmas by instituting the same requirements and consequences for everyone. Schwandt (2002) articulates an approach to evaluation that recognizes these moral dilemmas. He explains,

“Evaluation as education” is less like an applied science and more like a moral science. It is less concerned with the provision and mastery of technical and narrowly instrumental knowledge of effectiveness and goal attainment, and more concerned with the lived practice of making evaluative judgments in specific situations. (Schwandt, 2001, p. 233)

This lived practice “is fundamentally a matter of specific people in distinct settings at particular times struggling to reach an understanding (an interpretation) of the value of their actions (taken or about to be taken) in circumstances

that are always characterized by disagreement” (Schwandt, 2002, p. 55). He argues for using hermeneutics as a basis for understanding the social world. Using hermeneutics involves coming to an interpretation and judgment, such as whether a Catholic school provides quality education to its students, based on the staff, students, parents, archdiocese, church, and neighborhood. It relies on human interactions and dialogue, including deliberating, developing understanding, and using persuasion to reach agreement. It is an ethical, moral, political activity. He argues that this understanding is more closely aligned to how we engage in evaluation as an everyday activity. Rather than a formal, technical matter, evaluating and assessing can be viewed as activities that we do every day.

Even on a dull day, a school evaluates itself several times. Teachers listen for signs of learning; principals assess whether the new schedule is working; custodians check whether the fire extinguishers are still effective. There’s no way to escape evaluation. (Reck & Coreil, 1983, p. 1)

Is it possible that a characteristic of Catholic schools that contributes to their effectiveness is that evaluation and assessment is a morally engaged, everyday activity? If so, the introduction of high-stakes testing may interfere with these everyday activities. An alternative approach to stimulating educational reform may be to foster the use of everyday evaluation activities.

If Catholic schools do not use high-stakes testing, then how do Catholic schools ensure accountability? Darling-Hammond (1989) articulates additional conceptions of accountability, including political (e.g., election of school board members), legal (e.g., litigation), bureaucratic (e.g., rules and regulations), professional (e.g., teacher qualifications and continued professional development), and market (e.g., parents choosing schools). Some of these mechanisms already exist in Catholic schools. Political accountability occurs through the institutional structures of Catholic schools in relation to the Catholic Church, such as the authority to govern the school residing in a local congregation or the archdiocese. Such governing mechanisms in Catholic schools vary extensively. Market accountability is a major means of accountability in Catholic schools. For decades Catholic schools have been wrestling with declining enrollments. While market accountability may help raise awareness of the need for reform, it does not necessarily provide a direction for future reform. In contrast, professional accountability may provide a direction for reform. The NCEA and a number of Catholic universities, for example, provide professional development opportunities for teachers and administrators of Catholic schools. In comparison to public schools, legal

and bureaucratic accountability are less evident in Catholic schools. An additional source of accountability is also rooted in the notion of a calling or vocation that a teacher or administrator may experience. Vocation refers to being called by God to do one's everyday work as an act of demonstrating love to one's neighbor. Catholic educators have a strong tradition of viewing the profession of education as a response to a calling to serve God; thus, they are also accountable for their service to God. How might the Catholic community strengthen and build upon these notions of accountability rather than on viewing assessments as a means of accountability?

In addition to other means of understanding accountability, Catholic educators may also want to consider other approaches to evaluation and assessment that do not depend on student testing. Some evaluators have articulated ways of practicing evaluation that are more consistent with this everyday notion of evaluation, but are also done in an intentional, systematic manner. Evaluation approaches that are rooted in qualitative paradigms tend to be more closely aligned with everyday evaluation and assessment activities. These practices may provide alternatives within the Catholic educational community that aim to restore Catholic schools, while legitimizing the moral aspects and complexity of Catholic schools and fostering, rather than detracting from, the sense of community in Catholic schools.⁵ First, Stake (2004) uses case study methodology and description to represent the context of a school and multiple perspectives regarding salient issues in a school. From this approach, evaluation ought to be "responsive," meaning to "orientate to the experience of personally being there, feeling the activity, the tension, knowing the people and their values. It relies heavily on personal interaction" (p. 86). The intent in a case study is to describe richly a school, for example, giving the reader a vicarious experience, or an experience of being there themselves. Based on such a description, the reader is then able to interpret the quality of the schooling that takes place.

Second, Abma (2003) uses storytelling from multiple perspectives, such as those of students, teachers, administrators, and community members, as a means to enhance understanding of practices. She argues that we tell stories in order to make sense of our lives. For example, teachers tell stories about their students that help them articulate what students are capable of in the classroom, the life experiences students have outside of school, and so on. Sometimes these stories actually prevent teachers from fully understanding and reaching their students. As a result, Abma generates stories from multiple

5 These approaches to evaluation may not be beneficial for interfacing with the dominant culture regarding educational assessment. For that, Catholic schools may want to turn to standards-based reform approaches.

perspectives, for example, from teachers and students. She then shares the students' stories with the teachers and has the teachers respond to them, and vice versa. Such a process encourages teachers to challenge their assumptions about their students and generate new stories that reflect a better understanding of their students. Third, Kushner (2000) articulates ways of personalizing evaluation. He argues that rather than asking students about their school, for example, the inquirer can ask about their lives. The extent to which the school emerges as part of their life story is evidence of its impact on them. These represent three ways of doing evaluation that recognize the particularities of everyday life.

Conclusion

With the implementation of state assessment systems and NCLB legislation in public schools, a new wave of assessment and evaluation activities has unfolded in public schools in an effort to reform education. A perception exists in the Catholic school community that reform efforts are also critical in Catholic schools. It is also a time when Catholic schools are positioning themselves to acquire a greater amount of public funding. Given these circumstances, how will Catholic schools respond? Will Catholic schools also heighten the authority of assessment and evaluation? Will Catholic schools attach high stakes to assessment and evaluation results? Integrating standards-based reform efforts, using evaluation and assessment for improvement, emphasizing the evaluation and assessment of religious education, and focusing on an everyday understanding of how we evaluate and assess may be means for supporting educational reform efforts in Catholic schools. High-stakes testing as a means of providing accountability will not likely support reform efforts; rather, facilitating other mechanisms for accountability will be critical. In 1949, the developers of the *Criteria for Evaluating Catholic Elementary Schools* recognized the extent to which they could hold schools accountable for the formation of students, as well as the limitations of evaluation:

The primary reason for the existence of the Catholic elementary school is that it may produce the “supernatural man who thinks, judges, and acts constantly and consistently in accordance with right reason illuminated by the supernatural light of the example and teaching of Christ.” To this end, the Catholic school seeks to provide instruction and experiences that will help the child grow “physically fit, economically competent, socially responsible, culturally fine, and morally and spiritually perfect.”...In the measure to which the school assists the child in his total growth and development, its worth may be judged. The supernaturalization

of man is, of course, accomplished by his co-operation with Divine Grace. The school does not pretend to guarantee the individual's co-operation with grace, but it can and must guarantee to supply every natural and supernatural means at its disposal to facilitate the child's formation in Christ. The strong, silent action of the Holy Spirit in the soul of the child through the agency of the teacher remains forever beyond the merely human power of evaluation. (Catholic University of America Department of Education, 1949, pp. 18-19)

As Catholic educators wrestle with new agendas for assessment and evaluation as exemplified in public schools through high-stakes testing, let us remember the "merely human power of evaluation." Let us recognize the limited authority of evaluation and assessment in Catholic education. Let us utilize it for its value and potential "to supply every natural...means at (our) disposal to facilitate the child's formation."

References

- Abma, T. A. (2003). Learning by telling: Storytelling workshops as an organizational learning intervention. *Management Learning, 34*(2), 221-240.
- Abma, T. A., & Noordegraaf, M. (2003). Public managers amidst ambiguity: Toward a typology of evaluative practices in public management. *Evaluation, 9*(3), 285-306.
- Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 12-65). Thousand Oaks, CA: Sage.
- Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- Campbell, D. T. (1981). Introduction: Getting ready for the experimenting society. In L. Saxe & M. Fine (Eds.), *Social experiments: Methods for design and evaluation* (pp. 13-18). Beverly Hills, CA: Sage.
- Catholic University of America Department of Education. (1949). *Criteria for the evaluation of Catholic elementary schools*. Washington, DC: The Catholic University of America Press.
- Center for Research on Evaluation, Standards, and Student Testing. (1999). *CRESST glossary*. Los Angeles, CA: Author. Retrieved July 5, 2006, from <http://www.cse.ucla.edu/CRESST/products/glossary.html>
- Christensen, G. S., Cohodes, S., Fernandes, D., Klasik, D., Loss, D., & Segeritz, M. (2007). *Private school participants in programs under the No Child Left Behind Act and the Individuals with Disabilities Education Act: Private school and public school district perspectives*. Washington, DC: U.S. Department of Education.
- Convey, J. J. (1992). *Catholic schools make a difference: Twenty-five years of research*. Washington, DC: National Catholic Educational Association.
- Convey, J. J. (2002). *Catholic schools still make a difference: Ten years of research, 1991-2000*. Washington, DC: National Catholic Educational Association.
- Convey, J. J., & Thompson, A. D. (1999). *Weaving Christ's seamless garment: Assessment of Catholic religious education*. Washington, DC: National Catholic Educational Association. (ERIC Document Reproduction Service No. ED440016)
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397-418.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., et al. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.

- Darling-Hammond, L. (1989). Accountability for professional practice. *Teachers College Record*, 91(1), 59–80.
- Goertz, M. E., & Duffy, M. C. (2001, March). *Assessment and accountability systems in the 50 states, 1999-2000* (CPRE-RR-046). Philadelphia, PA: Consortium for Policy Research in Education. (ERIC Document Reproduction Service No. ED450639)
- Greene, J. C. (1999). The inequality of performance measurements. *Evaluation*, 5(2), 160–172.
- Gross, J. (2006, January 27). Florida court rules against voucher program. *National Catholic Reporter*, 42(3), p. 8.
- Haertel, E. H., & Herman, J. L. (2005). A historical perspective on validity arguments for accountability testing. In J. L. Herman & E. H. Haertel (Eds.), *Uses and misuses of data for educational accountability and improvement* (pp. 1–34). Malden, MA: Blackwell.
- Harris, J. C. (2000). The funding dilemma facing Catholic elementary and secondary schools. In J. Youniss & J. J. Convey (Eds.), *Catholic schools at the crossroads: Survival and transformation* (pp. 55–71). New York: Teachers College Press.
- Herman, J. L. (2007). *Accountability and assessment: Is public interest in K–12 education being served?* (CREST Report 728). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED499421)
- Joint Committee on Standards for Educational Evaluation. (1994). *Standards for evaluations of educational programs, projects, and materials* (2nd ed.). Thousand Oaks, CA: Sage.
- Kushner, S. (2000). *Personalizing evaluation*. Thousand Oaks, CA: Sage.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4–16.
- Linn, R. L. (2001). *The design and evaluation of educational assessment and accountability systems*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED455286)
- Linn, R. L. (2004). Accountability models. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 73–93). New York: Teachers College Press.
- Linn, R. L. (2006). *Educational accountability systems* (CSE Technical Report 687). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED492875)
- Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, 31(6), 3–16.
- Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco: Jossey-Bass.
- McAndrews, L. (2003). Recent history of Catholic schools and public funding, 1960 to present. In J. J. Augenstein, C. J. Kauffman, & R. J. Wister (Eds.), *One hundred years of Catholic education* (pp. 155–175). Washington, DC: National Catholic Educational Association.
- McLaughlin, T. H., O’Keefe, J., & O’Keefe, B. (1996). Setting the scene: Current realities and historical perspectives. In T. H. McLaughlin, J. O’Keefe, & B. O’Keefe (Eds.), *The contemporary Catholic school: Context, identity and diversity* (pp. 1–21). London: Falmer Press.
- National Catholic Educational Association Department of Elementary Schools. (1965). *Criteria for evaluation of Catholic elementary schools*. Washington, DC: The Catholic University of America Press.
- Nelson, M. S. (2000). Black Catholic schools in inner-city Chicago: Forging a path to the future. In J. Youniss & J. J. Convey (Eds.), *Catholic schools at the crossroads: Survival and transformation* (pp. 157–177). New York: Teachers College Press.
- Neuwien, R. A. (Ed.). (1966). *Catholic schools in action: The Notre Dame study of Catholic elementary and secondary schools in the United States*. Notre Dame, IN: University of Notre Dame Press.
- Notre Dame Task Force on Catholic Education. (2006). *Making God known, loved, and served: The future of Catholic primary and secondary schools in the United States*. Retrieved August 29, 2008, from http://iei.nd.edu/assets/4828/tf_cover.pdf
- Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice*, 15(3), 311–319.

- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks, CA: Sage.
- Powers, M. (1997). *The audit society: Rituals of verification*. Oxford, England: Oxford University Press.
- Reck, C., & Coreil, J. (1983). *School evaluation for the Catholic elementary school: An overview*. Washington, DC: National Catholic Educational Association.
- Reck, C., & Coreil, J. (1984). *Verifying the vision: A self-evaluation instrument for the Catholic elementary school*. Washington, DC: National Catholic Educational Association. (ERIC Document Reproduction Service No. ED251363)
- Robelen, E. W. (2008, May 6). Iowa mandates "core curriculum" for private schools. *Education Week*, 27(36), pp. 1, 6.
- Ryan, K. (2002). Shaping educational accountability systems. *The American Journal of Evaluation*, 23(4), 453–468.
- Schwandt, T. A. (2001). Understanding dialogue as practice. *Evaluation*, 7(2), 228–237.
- Schwandt, T. A. (2002). *Evaluation practice reconsidered*. New York: Peter Lang.
- Shadish, W. R., Jr., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories and practice*. Newbury Park, CA: Sage.
- Stake, R. E. (2004). *Standards-based and responsive evaluation*. Thousand Oaks, CA: Sage.
- Taymans, M. F. (1999). *Validating the vision: An assessment protocol for mission effectiveness, institutional accreditation, and strategic planning in the Catholic high school*. Washington, DC: National Catholic Educational Association. (ERIC Document Reproduction Service No. ED430301)
- Thomas, G., & Pring, R. (2004). *Evidence-based practice in education*. New York: Open University Press.
- Thompson, A. D., & Philibert, P. J. (1982). *That they may know you*. Washington, DC: National Catholic Educational Association.
- Trinder, L. (2000). Introduction: The context of evidence-based practice. In L. Trinder & S. Reynolds (Eds.), *Evidence-based practice: A critical appraisal* (pp. 1–16). Oxford, England: Blackwell Science.
- United States Conference of Catholic Bishops. (1998). *Protocol for assessing the conformity of catechetical materials with the catechism of the Catholic Church*. Washington, DC: Author.
- United States Conference of Catholic Bishops. (2005). *Renewing our commitment to Catholic elementary and secondary schools in the third millennium*. Washington, DC: Author.
- United States Conference of Catholic Bishops. (2008). *Doctrinal elements of a curriculum framework for the development of catechetical materials for young people of high school*. Washington, DC: Author.
- Weiss, C. H. (2004). Rooting for evaluation: A cliff notes version of my work. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 153–168). Thousand Oaks, CA: Sage.

Leanne M. Kallemeyn is an assistant professor in the School of Education at Loyola University Chicago. This paper would not have been possible without the thoughtful feedback and engagement from Dr. Lorraine Ozar and Dr. Michael Boyle from the Center for Catholic School Effectiveness at Loyola University Chicago; Dr. Ann Marie Ryan from Loyola University Chicago; Diane Raiche and Christina Gergits from the National Catholic Educational Association; and anonymous reviewers. Correspondence concerning this article should be sent to Leanne M. Kallemeyn, Ph.D., Loyola University Chicago, School of Education, 820 North Michigan Avenue, Chicago, IL 60048.