

Chapter 1

Waldorf Education and Education Reform

In this book we present research undertaken in order to document how teachers in independent Waldorf elementary schools set learning goals and evaluate their students' progress toward meeting those goals. The Waldorf approach to assessment contrasts strikingly with the public-policy view in the current debate on education reform, which favors assessment and accountability with norm-referenced high-stakes standardized testing. The view of Waldorf educators is that assessments that are designed for the purpose of supporting further learning, that are comprehensive of a range of abilities, and that are embedded in the daily classroom curriculum give a more holistic and therefore more accurate picture of a child's attainments. This approach is based on a particular conception of the stages of child development that informs the Waldorf curriculum. And it reflects a humanistic understanding of the goal of education that differs from public-policy formulations focused on economic success.

High-stakes standardized testing currently dominates the thinking of public-policy makers and the general public; it has become dogma that mandatory testing is the only possible way to assess students' learning, teachers' teaching, and schools' performance in relation to whatever standards are official in any given case. In contrast, none of the nine teachers who participated in the case studies reported here in Part II used a standardized or externally-made test during the five months that their teaching and assessment methods were documented, nor did such tests figure in their later plans. Instead they drew on a time-tested repertoire of classroom assessment strategies, most of which are comparable to the alternative methods now advocated by many professionals in the field of education.

The Relevance of the Waldorf Perspective

The researchers think that this study and its related discussions of accountability and teacher evaluation in independent Waldorf education will enable education professionals outside the Waldorf movement to better understand Waldorf

methods. Better understanding will make it possible for the Waldorf perspective to be included in the long-standing debate on how to improve educational outcomes.

The Waldorf perspective is relevant to discussions of several particular areas at the forefront of education reform, such as how to conduct more comprehensive and individualized assessment of capabilities important for learning beyond the measuring of math and reading skills; assessment literacy for teachers (i.e., upgrading their skills to include the conducting of alternative assessments); the importance of student-teacher and student-peer relationships for motivation and learning; a different understanding of the purposes of assessment; how to evaluate teachers in a way that leads to improved teaching rather than to rewards and penalties; and how to hold teachers and schools accountable for doing their job without external measures.

Waldorf education, founded by Rudolf Steiner in Germany in 1919, is independent and nonsectarian, offering a developmental (age-appropriate), experiential curriculum integrating academic subjects, practical activities, and the arts for students from preschool through twelfth grade. The teaching faculties in Waldorf schools that are accredited by the Association of Waldorf Schools of North America (AWSNA) are responsible for administration and all pedagogical matters including the setting of student learning standards; implementation of the Waldorf curriculum; frequency and formats of student assessments; mentoring and evaluation of teachers; and accountability to students, families, and the communities served by the schools.

The Waldorf movement is undergoing rapid growth: in North America today there are over 250 established and developing Waldorf schools and 14 Waldorf teacher education programs; worldwide there are some 900 schools in 83 countries. When he established the first Waldorf School for the children of factory workers in Germany in 1919, Steiner's view was that every child deserves a world-class education, the kind of education that at that time was available only to the wealthiest families. He wanted schooling to provide a pathway for every child to fulfill his individual potential and at the same time for education to be a source of social renewal.¹

¹Rudolf Steiner (1861-1925) was an Austrian-born writer, scientist, philosopher, artist,

Like some critics of the current fixation on accountability exclusively through standardized testing, Waldorf educators support Steiner's position that education is a cultural activity and as such must be independent of political and economic influence.² Accordingly, they maintain that standards, curriculum, accountability, and assessment should be in the hands of professional teachers who carry the day-to-day responsibility for the schools they teach in and the children they teach.

The fact that Waldorf schools are independent is not a reason to exclude the Waldorf perspective from the debate on public education reform. Given the urgency of finding solutions, a public-private partnership with exchange of best practices is sensible. Moreover, to judge from the scholarly literature on alternative assessment, many teaching methods and assessment strategies that are hallmarks of Waldorf pedagogy have been and continue to be used in non-Waldorf settings.³

As government-mandated standardized testing has escalated, many education professionals, among them respected scholars, have been proposing a variety of effective, feasible assessment alternatives with, however, negligible impact on public policy. According to Ravitch (2010), the heavy and inappropriate influence of the corporate business sector and certain large foundations on education panels and think-tanks is responsible for the current disconnection between the recommendations of educators and actual public policy with its promotion of "market-based reforms" (pp. 69, 200). What we report here can be taken as suggestive of what teachers who are free of that influence might be able to accomplish.

As we will show, the Waldorf approach accords well with recommendations coming from scholars in the fields of education and psychology aimed at improving

educator, social theorist, and spiritual researcher, who lectured and was well-known throughout Europe in the first two decades of the twentieth century. See <http://www.steinerbooks.org> for his publications available in English. See Appendix G and <http://www.wilnywaldorfschools.org> for the history of Waldorf education and Waldorf statistics for North America.

²See Appendix H for the Waldorf perspective on the proper but limited role of government in education, discussion of funding reforms that would help ensure real school choice, and Steiner's social theory as context for understanding Waldorf pedagogy.

³Several dozen public charter schools that have been established in the past decade are working with aspects of the Waldorf curriculum. Our work is focused only on independent Waldorf elementary schools that are recognized as full members of AWSNA; see discussion in chapter 10.

educational outcomes. The Waldorf experience validates the direction of many of these recommendations, for example, advocacy for more comprehensive (beyond cognitive skills) assessments and greater reliance on performance rather than written test assessments. Waldorf teachers are trained to design and implement these more comprehensive kinds of assessments. They know how to effectively integrate strategies like group projects, student self-assessment, and artistic work into their teaching. They understand the importance of a healthy social classroom dynamic that supports learning and the pedagogical benefits of a close teacher-student personal relationship based on mutual trust.

What is the Purpose of Education?

The direction taken by educators seeking appropriate and effective ways to measure student learning—the central concern in today's debate—will depend, to begin with, on what they understand as the purpose of education. Whom are we educating and what is an education for? What are we assessing and why? Whoever works with these questions must confront the seemingly conflicting needs of the individual and the society. That is, every human being is unique but at the same time participates in communities, societies, and nowadays a world requiring common knowledge, skills, understandings, and values. No two children are exactly alike and each has specific individual needs and talents; at the same time all children, whether born in Uganda or America, go through similar stages in their development. All deserve an education enabling them to fully realize not only their individual potential but also the goal of becoming useful citizens in a shared world.

We assume that most educators would agree, therefore, that the goal of education is twofold at its core: to bring out the best in the individual and help the individual become a citizen in society, to build a common understanding and shared base of knowledge while at the same time sparking individual aspirations for its students. Education must meet these two criteria simultaneously in order to be successful. We would say that, given this task, educators must find a balance between what in the curriculum needs to be standard and what must be flexible to allow for creative teaching, between teaching that addresses common goals and teaching that addresses the needs and talents of the individual.

However, despite any such agreement, assumed or real, the two disparate approaches noted above for improving educational outcomes dominate today's debate: the public-policy approach focused on testing, on the one hand, and diverse proposals by education scholars for alternatives to testing, on the other.

Both approaches have roots in the history of education reform. On the one hand, a long-standing trend toward administrative efficiency has led to a one-size-fits-all approach to educating students. On the other hand are those education scholars, going back to Dewey (1938/1997), who argue that richer curricula, experiential learning, and multidimensional individualized assessment rather than testing will result in better learning outcomes. Both these approaches have helped to shape the discussion, policy, and practice of education in the United States since the beginning of the twentieth century.

Public Education and the Social Efficiency Model

Administrative progressives in the early part of the twentieth century laid the ground for today's public-policy approach. Educators like Edward Thorndike and Ellwood Cubberley wanted to make education efficient and create "tracks" that directed young people toward a place in society based on IQ scores and other tests. Labaree (2010) calls this the "social efficiency" model of education wherein young people are positioned in society by virtue of their birth, talents, and social standing (p. 94).

Another aspect of the social efficiency model is its view of learning and assessment as two separate activities: first, students acquire information and skills; later, teachers or some outside authority make assessments to measure that acquisition. Shepard (2000) argues that this model fails to support good learning outcomes. She traces the historical development of this paradigm (acquisition of knowledge followed by assessment of acquisition) through its manifestations in behaviorism, theories about social efficiency, and scientific measurement, a development culminating today in bias among educators at all levels against "subjective assessments" and powerful beliefs held by teachers about the need for objectivity attainable only by separate assessment of specific instructional goals (pp. 4-6). In contrast to theories about knowledge acquisition as the accumulation of information, her framework acknowledges new insights from cognition theory,

especially "that learning is an active process of mental construction and sense making" (p. 6).

Education influenced by the social efficiency model has been likened to assembly-line production by Sir Ken Robinson (2007) in his numerous TED talks on creativity in education. He argues that the one-size-fits-all approach that is typical of schools fails to address the creative potential that each student has to offer. We are sketching here the broad trends and in fairness must note that many teachers inside their classrooms resist today's assembly-line prescriptions for teaching. Many teachers and principals are using their creativity to work within a system that could otherwise be characterized as stultifying.

In public education in the United States, the social efficiency approach has evolved into today's reliance on standardized tests. The newest developments in federal- and state-level education policy confirm that standardized testing has captured the imagination of public-policy makers and the general public. Every step forward in public policy entails more testing. The federal response to diminishing returns from testing mandated by the No Child Left Behind (NCLB) legislation, for example, was a huge funding authorization for more sophisticated standardized tests, to be given more frequently (Dillon, 2010; Duncan, 2012; Hu, 2011; U.S. Department of Education, 2010b, 2012; see also FairTest 2013).

The public perception that a top-down model of imposed standardized testing is the only effective way to measure students' learning and evaluate teachers' performance seems unshakable. This perception is continually reinforced by the work of think-tanks, foundations that are focused on education reform, and pronouncements and initiatives by policy makers at all levels (e.g., Ravitch, 2010, pp. 200-219; U.S. Department of Education, 2010a, 2010b, 2012).

The Scholars' Critique of Testing

Many education professionals question the trustworthiness of standardized tests, pointing out their narrow focus and consequent limited usefulness for assessing the full scope of a student's capacities and knowledge or for pointing to individual gaps in understanding. Among them are Bassett (2005), Carey (2008), Nichols and Berliner (2007), Ravitch (2010), Shepard (2000), Stiggins (2002), and Wiggins (1990, 1998). Finding effective alternatives to standardized testing of students

is critical because test results are used not only for measuring student learning but also as a basis of rewards and penalties for teachers and schools. Though they can be useful for identifying problems and gaps, such tests are widely recognized as unfair, unreliable, and counterproductive of learning (e.g., Ravitch, 2010, pp. 152-154).

Some scholars have challenged standardized tests from the standpoint of equity, calling attention to biases that make their results invalid, unreliable, or unfair for measuring abilities of, especially, minority students (e.g., Ayrasian, 2005, pp. 113-115; Au, 2009; Estrin, 1993; Hilliard and Amankwata, 2003; Rothman, 1994). Shepard (2000) notes the well-known "pervasive negative effects of ... externally imposed testing programs" on teaching and learning, including indications of decreases in student understanding and, of greater concern, of "de-skilling and de-professionalization of teachers" (p. 9).

High-stakes standardized tests are rarely educative, that is, supportive of further learning, because they provide no feedback to students (Wiggins, 1998). Tests are poor at measuring understanding, which itself comprises a range of cognitive and affective abilities (Darling-Hammond et al., 2008; Wiggins & McTighe, 2005). Nonetheless, some scholars advocating alternative methods of assessment assume that non-test assessments still need to be confirmed by standardized norm-referenced tests made by subject experts, for example, those provided by textbook publishers or computer modules in math and reading. Most also accept additional testing as mandated by federal legislation.

The position of Waldorf educators (Mitchell, Gervin, Schubert, Mancini, & Hofrichter, 2008) is unequivocally against all standardized testing at the elementary-school level, for reasons that will become clear as we proceed. They make the point, among others, that standardized tests cannot measure development of individual character traits and dispositions critical for learning and retention, such as diligence, interest and enthusiasm, initiative and resolve, creative and independent thinking, and collaborative skills.

The Search for Alternative Methods of Assessment

Disillusionment with standardized testing has accelerated a long-standing and widespread search for better methods of assessing student learning, led by

scholars continuing the second approach noted above. Education professionals and even some public-policy makers are calling for "alternative" assessments, defined by Carey (2008) as anything other than tests and usually including performance assessments, problem-based assessments, student self-assessment, and various forms of classroom assessment aimed at supporting multiple and individual learning styles.⁴

An extensive and growing body of scholarly literature endorses well-planned instructional (i.e., classroom) assessments as most likely to provide the fairest, most valid, and most reliable assessments (e.g., Anderson, 2003; Carey, 2008; Goodman & Carey, 2004; McMillan, 2006; Shepard, 2000; Stiggins, 2001, 2002). And a large and growing number of studies explore the effectiveness of performance assessments, that is, assessments conducted while students are engaged in a learning task, as tools for learning as well for assessment *per se* (e.g., Brualdi, 1998a; Doane, Rice, & Zachos, 2006; Marzano & Kendall, 2007; McTighe & Ferrara, 2011; Mehrens, 1992; Moskal, 2003; Wiggins, 1998; Wolf & Reardon, 1996).

With calls for "authentic" assessment (Tombari & Borich, 1999; Wiggins, 1990), "ubiquitous" assessment (Goodman & Carey, 2004), and more "coherent" (Herman, 2010) approaches to assessment, educators acknowledge the importance of more comprehensive assessments that would include capabilities like problem-solving and collaboration and would assess literacy and competence in other subjects besides math and reading (the focus of most standardized tests).

The legislative agenda of the educational leadership organization ASCD, whose membership is mainly teachers, urges education reform to meet the needs of the "whole child" and assessment that provides a "more comprehensive picture of student achievement" through multiple measures (ASCD, 2012).

Outside of academia, some organizations are heralding methods of assessment that entail sophisticated but still standardized measures, along with increased use of technology in the classroom. Edutopia, the website of the George

Lucas Educational Foundation, advocates performance-based, comprehensive assessment of the "whole student" with games, team projects, class presentations, portfolios, and other activities; it claims an important role for computers and other technology in classrooms and better standardized tests for assessment.⁵

Another example is Partnership for 21st Century Skills, which promotes tools and other resources, including high-quality standardized tests, for measuring compliance with common state standards. Their model assumes and endorses centralized planning with state-mandated standards and state implementation of these assessments.⁶ We look more closely at P21 in chapter 7.

The educative assessments advocated by Wiggins (1998) and Wiggins and McTighe (2005) are curriculum-embedded assessments that give students personalized, individual feedback based on what was taught in the classroom, how well the students understand the concepts, and where they are on a continuum of mastery. This also gives the teacher immediate feedback, which can help her modify the lessons as needed.

Other educators cite the need for models (real cases) of effective learning-centered, curriculum-embedded, more individualized and comprehensive assessments (e.g., Black, Harrison, Lee, Marshall & William, 2004; Moskal, 2003; Shepard, 2000). It is the need for examples of these kinds of alternative assessments in practice that makes the inclusion of Waldorf methods in public and scholarly discussions important, or so we think.

The first Waldorf school in the United States was the Rudolf Steiner School founded in New York City in 1928. Thus, Waldorf education has a track record that has run parallel to the evolving policies and practices of mainstream education for almost 100 years. Waldorf educators have always stood firmly on the side of those advocating for richer curricula, multidimensional, individualized, and curriculum-embedded assessment. Many of the alternative assessment strategies recommended today in the scholarly literature can be observed in Waldorf classrooms, where they have been used for decades.

⁴The focus of much of this literature is teachers' assessment literacy: teaching teachers the skills they need for making competent alternative (to tests) assessment (e.g., Anderson, 2003; Arter, 1999; Heifner, 2004; O'Leary, 2008; Popkum, 2009; Stiggins, 2002).

⁵See <http://www.edutopia.org>

⁶See <http://www.p21.org>

Waldorf education also holds firmly to its foundational understanding of the purpose of schooling, which we discuss in chapter 2 and which—to use phrasing often found in the outreach materials of Waldorf schools—at its heart is a commitment to support students in developing into free-thinking human beings who can impart meaning and direction to their lives.

The Plan of the Book

The next chapter (chapter 2) introduces the pedagogical foundations and principles of Waldorf education, which provide the rationale for teachers' methods of student assessment. We begin necessarily with the answer that Waldorf educators give to the question, *What is the purpose of education?* The chapter continues with a summary treatment of Waldorf pedagogical fundamentals for elementary-level teaching, evident in all the teaching documented in Part II.

Part II (chapters 3 through 7) reports our work with nine Waldorf teachers, whose teaching and assessing we followed for five months. Chapter 3 describes our research methods. The next two chapters report how the teachers set learning goals and planned teaching (chapter 4) and their methods of student assessment (chapter 5). A separate chapter (chapter 6) discusses formal reports, both those that the teachers in the case studies made and the practice of formal reporting in Waldorf education generally; it also touches briefly on the official student learning standards that were in place in the schools that participated in the case studies.

In chapter 7 we discuss Waldorf teaching and assessment methods—those we documented as well as those we understand as ideal Waldorf practice—in relation to the work of selected education scholars. We focus on scholars whose recommendations are comparable or convergent with the Waldorf approach. At the same time we highlight the Waldorf concept of age-appropriateness, which we think is unique and which is based on a particular understanding of the developmental path of the growing child.

We devote Part III (chapters 8 through 10) to the topic of accountability. We address the questions: *Why should we trust Waldorf teachers? Are their assessments valid and reliable in the absence of external measures? How can Waldorf schools be held accountable in the absence of externally-applied standards and government oversight?* To answer, we offer a full account of the internal mechanisms in Waldorf

education aimed at ensuring the quality of education provided by AWSNA-member schools. These include AWSNA's path to membership for schools and teacher training institutes and the professional preparation of Waldorf teachers (chapter 8) and three examples of AWSNA-approved teacher mentoring and peer evaluation systems (chapter 9); the latter are shown to be comparable to teacher evaluation systems that are proving effective in sectors of public education. Teacher evaluation has emerged as a central concern in education reform; this account of Waldorf practice should be a welcome contribution to discussions of this topic.

Chapter 10 offers other evidence to justify confidence in the validity of Waldorf teachers' student assessments, with the caveat that a requirement by critics for the kind of proof provided by standardized-test scores is incompatible with the Waldorf understanding of the function of assessment. We recapitulate our findings on Waldorf assessments, emphasize that the Waldorf understanding of age-appropriateness is not the same as what non-Waldorf teachers mean when they say "at grade level," and discuss school independence as a defining condition for a school to be a Waldorf school. We end by suggesting a series of questions that could profitably be raised in light of an informed understanding of Waldorf education and its methods of assessment.