Assessment: A Waldorf Perspective

Martyn Rawson

According to the Swiss Waldorf teacher Robert Thomas, "for as long as schools have existed…one of the teachers' main tasks has been to observe, evaluate, judge and classify the work of their students." (Thomas 2005) Thomas goes on to point out that the question of assessment has three central aspects: the *what*, the *how*, and the *who*. The 'what' consists of observable, measurable facts; the 'how' refers to the relationship between the learner and the teacher; and the 'who' indicates the person learning and developing. What we understand about a person is something unique; it cannot be standardized, generalized, or measured. It is also never complete, but always remains open.

We generally assess the past, what has already happened. But assessment also means getting a sense of what is emerging, what is in a state of becoming. We can try to understand the future of the child, as it emerges, to borrow Otto Scharmer's phrase. (Scharmer 2009) An assessment can also be a kind of message from the future that we, as teachers, attempt to read. This means that we have to create space in our assessment for the person's potential development, and doing this well can even help a healthy future come about. Therefore, in assessment we need to give space to the voice of the person we are assessing and be open to hints of what he or she may become—the distinctive signature of that person as it manifests in deeds, gestures, and acts of creation. In doing so, we must be most careful not to limit that potential through our judgments based on the past.

We have to create space in our assessment for the person's potential development. … In assessment we need to give space to the voice of the person we are assessing.

The *who* refers to the person doing the assessing. How we assess sends a message and defines the nature of that relationship. Is it well-intentioned, respectful, and caring, or distancing, objectifying (i.e., making the person into an object, or even a statistic), or labelling (weak, average student or good student)? The quality of an assessment depends on the quality of seeing, listening, and understanding.

Assessing learning in a Waldorf context assumes we know what learning means from an anthroposophical perspective. There has been surprisingly little published on what the Waldorf view of learning actually is, although Jost Schieren recently made an interesting attempt. (Schieren 2012) Learning is a complex theme that I can't go into here, but perhaps the most important things to remember are that learning is a process that transforms the whole person and thus changes the way we are and how we act, and that it has to do with making and sharing meaning about the self, others, and the world.

Let me state at the outset my position regarding assessment for learning. I was jointly responsible for developing and publishing probably the first public definition of learning outcomes grade for grade (for math, English, and first foreign language) in the Waldorf literature. (Mepham & Rawson 1997; Rawson & Richter 2000) Although I have regularly promoted the use of assessment for learning and criticized the lack of good assessment practices in Waldorf schools, over the years I have become concerned that the pendulum has swung too far the other way.

This article revisits one of my basic interests in Waldorf pedagogical principles and is a much shortened version of a longer essay on this topic.

The context of assessment today

Assessment today has to be understood within the context of the current dominant educational culture of emphasizing performance, surveillance, control and accountability using instruments of standardization, testing, and ranking. This culture is so pervasive and accepted by so-called common sense (who can argue against what is called "quality in education"?) that there appears to be no alternative. However, Professor Stephen Ball of the London Institute of Education dubs what he calls performativity in education "a culture or system of terror. It is a regime of accountability that employs judgments, comparisons, and displays as means of control, attrition, and change. The performances of individual subjects or organizations serve as measures of productivity or output, displays of quality, or moments of promotion or inspection." (Ball, 2008, 49) Ball has explored the worldwide networks that promote such educational policies in search of profit without a grain of pedagogy in their systems. (Ball, 2012) This is not conspiracy theory, but serious academic work.

The impact of such policies has also been well researched. Professor Paul Verhaeghe has identified that an educational culture that emphasizes competition—reinforced by selection and grades and driven by the motivation of acquiring educational capital in the form of ever more inflationary qualifications that say nothing about a person's human qualities—leads to a "winner/loser" mentality among students. (Verhaeghe 2015) Those who buy into this system tend to value materialism, consumerism, individualism, narcissism, intolerance of difference, xenophobia, and the belief that other people's misfortune (e.g., unemployment, sickness, burnout, lack of health insurance, refugee status, etc.) is their own fault. Even among the "winners" in the education "market" there is a high frequency of depression, insecurity, substance abuse, and emotional exhaustion from the relentless pursuit of success and bodily youthfulness. Many intelligent, highly-educated, and successful still-young people reach a point in their lives when burnout makes itself apparent and the long, difficult search for meaning and holistic health starts. Verhaeghe asks whether it might be wiser for such societies to offer a different kind of education, one based not on competition but on social solidarity.

> **What we understand about a person is something unique; it cannot be standardized, generalized, or measured.**

This educational treadmill starts rolling in preschool. Ulrike Keller has researched how grades affect young children's learning behavior. (Keller, 2012) Those who can afford it place their children as early as possible in "turbo-kindergartens" that prepare children for school and give them an edge over their "competitors." Once at school the competition begins in earnest. Keller compares the different effects on a child when she comes home and the first question is, "What grade did you get in school today?" as opposed to, "What interesting things did you experience in school today?" The first question doesn't even have to lead to tangible rewards (something to consume or just the money) to make it clear to the child which priorities the adult world has.

This leads us back to the question of what we understand and value in terms of learning. What is learning actually about? Professor Horst Rumpf has described two ways of learning about the world and the knowledge that these produce. (Rumpf, 2010) The first involves capturing, dissecting, analyzing, labelling, categorizing then mastering and exploiting aspects of the world. (The choice of verbs already describes

the attitude behind this approach.) The second way of learning involves respectful and careful watching (in Waldorf we would say, "with wonder") in a mood of gratitude for being able to share this experience, combined with the assumption that there is always more than we can perceive or understand at any one moment. It is a descriptive process and its purpose is to enable us to become more richly experienced and thus to transform ourselves as learners.

The first method of learning can be assessed by testing whether the learner *possesses* the knowledge. Here what counts is the outcome, and we can measure it quantitatively. In the second type of learning it's the process that counts. We can assess qualitatively how the knowledge was gained and how the person has changed through this rich experience. In Waldorf education we need both types of learning and both types of assessment. The difference lies in which learning method we teach first and which one we privilege and reward.

Assessment is moral action

Assessment always involves making judgments about what we value, about how we recognize quality, excellence, effort, engagement, intentions, and the effects our pupils have on their surroundings. How we value and reward performance thus has a moral message. This permeates the pedagogical culture in the school. Assessing only cognitive academic achievement lowers the value of everything else children do. The attainments we recognize, the awards we grant, how we praise and celebrate—all this has a moral effect that shapes the young person and also says a lot about us. Therefore, we need to be very clear about our values.

> **We value people who learn from mistakes rather than those who claim that to succeed means never to fail.**

> **Assessing only cognitive academic achievement lowers the value of everything else the children do.**

Of course, Waldorf schools need to come to terms with external social expectations regarding the measurement of attainment so as not to isolate the pupils from their community. Our students must get the necessary qualifications for learning a profession or going to university. Waldorf education does indeed value and promote personal achievement, while at the same time espousing a non-competitive approach to learning. We value, for example, individual effort regardless of the actual outcome. We value the way people are and not just what they say. We value people who learn from mistakes rather than those who claim that to succeed means never to fail.

The question remains, however, whether we adopt commonly-held practices that conflict with our own values, such as an emphasis on grades and cognitive achievement and competitiveness. How soon and in what way do we encourage our students to learn for the test or exam? For me the central question about assessment is: What effect does it have on the whole child's being, including her future? Does it unlock doors, open windows, assist movement (inner or outer), and encourage learning without determining its outcome?

Assessment within a developmental approach

A central tenet of Waldorf teacher education is that teachers should understand not only human learning but also human development from an anthroposophical perspective. Careful observation and understanding of each individual leads us back to an appreciation of the general and common processes of learning and development. However, it is important to see these general principles not as norms but as archetypal processes that come to individual expression differently in each person.

Knowing this enables teachers to design and plan their own lessons. It is an iterative process of doing, observing, contemplating, and reflecting that informs the creation and shaping of pedagogical situations in which learning and development can occur. This also means that we need to accommodate our modes of assessment to the developmental situation of the child.

Here there is a risk that we impose inappropriate developmental frames onto children, by, for example, having expectations of what a child in grade three should be and be able to do. We know from the work of the Swiss pediatrician Remo Largo, who gave a keynote lecture at the Goetheanum at the most recent World Teachers' Conference, that the average *normal* developmental span at the age of 6 is three years and at the age of 14 it is five years (not including children with special developmental issues). (Largo 2011) Therefore there are no standard grade 1 or grade 8 students. Each child has a unique developmental biography and will therefore respond to the same experiences in different ways. That is normal.

The Waldorf curriculum is structured into years/grades, and the subjects follow their own developmental logic with regard to content and method of teaching. The wisdom of this curriculum is that it suggests a sequence of experiences that can have a harmonizing effect on the children. (Rawson & Richter 2000) It offers developmental tasks to a heterogenous group of children of the same chronological age who comprise a community of learners through their shared experiences. This, however, does not mean that outcomes have to be standardized.

To be inclusive as a teacher, I need to differentiate the tasks I give children who have had the same experience or heard the same story. Some children need a task that simply

...[T]he most important factor in enhancing learning is teachers who are able to observe the effects of their teaching on their pupils and who can modify their approach to teaching based on these observations.

requires them to repeat in their own words what the teacher presented; others can be asked to make comparisons, interpret, modify, or even find their own examples—in other words, to learn in a transformational way. Children process their experiences differently because they are different. Therefore, if I design a pencil and paper test, I have to make it possible for all learners within the class to participate successfully, including those whose development for some reason lies outside the broad normal span of development. It should be possible for everyone to succeed at some level. That means I expect more from some than from others. There are, of course, better ways to assess children than by giving them pencil and paper tests.

Assessment of teaching comes before assessment of the pupils' performance

Assessing pupils assumes that the teaching affords optimal learning. If that is the case, then the only variable factor is the pupil. But unless the teaching is optimal, the focus needs to be on the effect rather than on the main variable in the cause—namely on the teacher, the teaching, and the learning context. Of course, the state looks at testing outcomes as a judgment on the quality of education in a school. But this is a crude instrument that does not take the context into account. The poor rating of schools in socially deprived areas always reflects unfairly on the teachers, who may in fact be doing a great job under almost impossible conditions.

A balanced approach has to look at both the teaching and the outcome of the teaching. Put differently, assessment involves a cycle of reflecting *on, in*, and *for* teaching. It means reflecting on pupils, their work, what they say and do, and how they do it. It involves knowing what to do in the moment in the classroom, and it means weighing up the next

move. *Reflecting on* is retrospective, looking back at what happened. *Reflecting in* or while teaching involves noticing how children respond and intuitively knowing how to respond pedagogically. It is a kind of pedagogical intuition that draws on embodied, tacit working knowledge acquired through practice, study, and contemplation. *Reflecting for* teaching involves planning, anticipating, and preparing lessons.

Assessment rarely requires a final or ultimate judgment. It is careful, respectful and tentative, provisional and iterative—in other words, it is ongoing. Assessment should be a way of supporting learning and development.

Research by John Hattie shows that the most influential factor in enhancing learning is teachers who are able to observe the effects of their teaching on the learning behavior of their pupils and who can modify their teaching approach based on these observations. (Hattie, 2012) Hattie's core message is: Know thy impact. We can translate this phrase to mean: "Perceive how children respond to the learning situations you create through your teaching, curriculum, school environment, and climate. Learn from this and then adjust your approach accordingly." This is the same as saying: "Base your teaching on practice-based research." This means focusing assessment on the teaching rather than the pupils by observing its effects on the learning.

The second most important factor, according to Hattie, is the nature of the teacher's response to the pupils' learning. The purpose of assessment for learning is to give the students reflections and suggestions in a form and in such a way that the pupils understand what the next concrete step needs to be. Supportive, tactful, helpful, sensitive, and practical responses are invaluable in learning.

Ipsative assessment looks at the pupil's biography, sensing what is emerging in the person over time and across different situations. It is certainly the most relevant way to assess individuals in classes with a wide range of abilities.

What kinds of student assessment are there in Waldorf education?

Assessment in Waldorf schools appears in two basic forms: assessment *for* learning and assessment *of* learning. Both methods can be used to monitor the quality of the teaching and to measure pupils' attainments against relevant criteria. The Waldorf approach is based primarily on assessment for learning and ipsative assessment, which evaluates the pupil in relation to her own personal starting point. (Avison and Rawson, 2014, 38) I emphasize that I am writing from a European perspective. I have no experience of Waldorf schools in the U.S. and some of my terminology may be unfamiliar. I will now outline the types of assessment that I think are relevant in any Waldorf context.

Ipsative-referenced assessment

Ipsative assessment (from *ipse*, the Latin word for "self") is self-referenced. That means a pupil's performance is evaluated against her own prior performance as judged from a given starting point (e.g., the start of a school year, the beginning of a module or block). In other words, the same achievement might mean a great improvement and effort for one person, or the result of little effort and no improvement for another. Ipsative assessment is a way of individualizing assessment to the person, taking that person's whole situation into account.

Ipsative assessment looks at the pupil's biography and senses what is emerging in the person over time and across different situations. In effect, this is the Waldorf way of assessing, and it is certainly the most relevant way to assess individuals in classes with a wide range of abilities, even including children with learning disabilities. It means the student is competing against no one but herself. The pupil asks, "Can I

do better than last time? Can I improve on what I have achieved so far?" " What do I need to do to be better at this?" Ipsative assessment gives the individual pupil answers to these questions.

An example of this form of assessment in my school is the use of response sheets after each main lesson block (or half-term or project) from ninth grade onwards. The students gather these sheets in a folder. In a tutorial conversation the student and the class advisor look back over the sheets and discuss trends and tendencies. Here the comparison is with the person herself, the difference between earlier and present achievements in the same field or subject, rather than comparisons with anyone else or against standardized criteria. The class teacher can do this when she gives the main lesson books back at the end of each main lesson block. The assessment is addressed to the student and is not measured in relation to the class norm or what the teacher thinks is suitable for this grade.

Formative assessment

Formative assessment is process-orientated. It includes the monitoring of the ongoing learning of individuals and groups, and it is used to make adjustments to the learning process and to offer reflections to students about their learning. It is concerned with how learning occurs and is either informally or formally noted by the teacher, either verbally or in writing. The documentation of performance observed *in situ* is actually quite difficult to manage since it tends to take one's attention away from the lesson. However, when we do it after the lesson, we are liable to the risks of selective memory. Sometimes I give visiting students a checklist on which they record what individual pupils do. If I compare these more objective observations with my own retrospective feelings, I will recognize when I am wide of the mark. Therefore I make a very short list of

important aspects I wish to observe in the lesson and mark their occurrence with a simple code.

Summative (outcome) assessment

In its original meaning, summative assessment means making judgments about whether outcomes have been achieved. In mainstream education it frequently means measuring performance against standardized and statistical criteria. However, in Waldorf education it means summing up what has been done. Perhaps a less-contested term would be "outcome" assessment. This kind of assessment can be used to establish whether a student has completed a task successfully or has learned what was required. It usually occurs at the end of a period or block of learning before a new stage or phase starts. It is perhaps most appropriately applied to assessing a literal performance (on stage in drama or eurythmy, musical performance, art exhibitions, giving a speech, demonstrating a practical task, and so forth).

Summative assessment should be carefully designed to be realistic and relevant, involve authentic problems or challenges, and provide an opportunity for creativity.

Outcome assessment is often formalized by tests with clear criteria for achievement. It requires a judgment based on evidence as to whether students have reached what they set out to achieve. This assumes a baseline of criteria (sometimes called learning outcomes), defining what the students are supposed to achieve and that the students know what is expected of them.

Outcome assessment sets tasks or uses criteria to ascertain levels of attainment. For example, at the end of a tenth grade biology class, the intended outcome could be: The student has understood the human digestive system. We ascertain whether she has or hasn't by asking her to draw a diagram and explain the digestive process using appropriate terminology. We thus establish how well she has memorized the material. However, this tests only one dimension of learning. We could then ask her to

compare the character and function of various organ systems and even ask how learning about these organs has affected her understanding of her own behavior.

In my view, unless we think it is important to test short-term memory (because learning facts by heart is usually short-term!), the students need access to all the relevant information they might need to answer the questions. If they have their notes and main lesson books at hand during the test, then the questions can aim for understanding rather than simple memorization. Since there is so much information available at our fingertips, the real skill is not remembering it but knowing how to use it. This approach to assessment encourages the students to keep good notes. Main lesson books should not just be beautiful and completed on time; they should above all be useful resources for the future.

There also needs to be a range of tasks since no one task is likely to involve all the skills, knowledge, and attitudes being assessed. These tasks can be collected in a portfolio over a period of time. The criteria for grading the students' work need to be clear to all involved and broken down into concrete steps or stages. In situations where holistic assessment is used—for example, judging creative writing and the quality of writing in essays—there needs to be a balance of hard (quantitative) and soft (qualitative) criteria.

Outcome assessment can also establish developmental markers, such as school readiness, reading and writing abilities, and levels of numeracy. This kind of assessment has a place in Waldorf education, but not as an instrument of selection, but simply of diagnosis.

Finally, having tasks assessed by teachers who don't know the student (external assessors) or who are not currently teaching them is useful in the dialogue among the teachers. Teachers can get a perspective on whether the tasks

were appropriate and whether the criteria were properly applied. This is of course not something the students see. Outcome assessment should always take other recent work into account because test situations often put students under pressure and may not reflect what they are capable of doing.

Continuous assessment

This method is not actually as continuous as it sounds. It is in fact punctuated and based on accumulating the outcomes of regular written assignments. In effect it is a form of outcome assessment built up over time rather than in a single test at the end of a period of learning. It could be used formatively by making regular observations of and documenting oral contributions, participation in group work and behavior, but usually it isn't. Ciborski and Ireland describe this as a method used in Waldorf schools, though it is very much a state school method. (Ciborski & Ireland, 2015, 89-90) Students report that continuous assessment raises their stress levels and can have the same emotional effect as continuous surveillance. Continuous assessment sounds rational and effective but, then again, who wants to be continuously assessed?

Diagnostic assessment

This can be used to identify whether individual students need support. Diagnostic assessment is sometimes associated with a focus on problems and deficits but doesn't need to be. One can identify examples of good learning and healthy development and learn a lot from these.

Assessment for learning

This approach seeks to interpret evidence so that learners and their teachers can determine where the learners are in their learning, where they need to go, and how best to get there. It involves sharing learning goals with pupils,

> **Outcome assessment… has a place in Waldorf education, but not as an instrument of selection, but simply of diagnosis.**

engaging them in self-assessment, and providing them with the teachers' perspectives. This process leads the students to the next step and assumes that every student can improve. (Hipkins 2007; Swaffield 2008; Swaffield & Dudley 2011)

Evaluative assessment

This is used by teachers to evaluate whether methods of teaching were effective. Here it is the teaching that is evaluated, not the students.

Peer-assessment and self-assessment

Peer assessment means that people are assessed by their fellow students. It also entails that students get involved in designing assessment procedures, which helps them to appreciate the nature of assessment and to conduct it in a considerate, fair, sensitive, and tactful yet honest manner. This process strengthens their own ability to assess themselves. It is a human and non-bureaucratic way of valuing the uniqueness of a person; in short, it provides the antidote to standardization.

Waldorf education values and promotes personal achievement, while at the same time espousing a non-competitive approach to learning. We value individual effort regardless of the actual outcome.

A complete assessment would involve a self-assessment, peer-assessment, and assessment by the teacher—for instance, on the occasion of eighth grade or twelfth grade projects. These three dimensions, which encompass the social context in which the learner is embedded, enrich, as Robert Thomas suggests, the cultural climate of the school.

Some essential things to consider before assessing

The first point to make about any assessment is whether the criteria we use are fit for the purpose. I advise being critical. Let me illustrate this with a somewhat extreme example. If I read in some official Waldorf catalogue of standards (e.g., as listed in Ciborski and Ireland in the Appendix) that a child of a certain age should

be able to sit on a chair, pay attention, and be tolerant of others (among other things) by the end of first grade, I need to consider a number of things before I make this assessment. For instance, what happens to the child if I observe him and then document his behavior? Does my observation affect the child? Steiner and contemporary social researchers, though for different reasons, insist that it does.

Secondly, what do I do with my assessment once I've made and documented it? What consequences does it have for the child? What can the child do and what are the parents supposed to do? Does my assessment change my practice so that the child can sit longer, pay more attention, or be nicer to her classmates, or is she supposed to effect this change simply because I ask (or tell) her? In other words, is this standard actually appropriate for this (or indeed any) child? The point is to ask: What effect does an assessment schedule actually have on the children being assessed, and what are the pedagogical values underlying this standard?

The example of a child sitting on a chair illustrates—perhaps in an extreme way—the point about our values in classroom management. I know Waldorf classes that are models of good behavior when the class teacher is present, but as soon as this authority figure is out of the room, there is chaos or bullying, or they become unteachable in the presence of subject teachers. I also know "good" classes who can work for longer periods (even in grade 2 or 3) when the teacher is out of the room and who are interested in and welcoming to subject teachers. My point is that the bad behavior of the children in the first example would be noted in their assessments without reference to the invisible class teacher effect.

Assessing skills, capabilities, and competencies

Much learning is social in that it occurs in dialogue or interaction with other people. Furthermore, most skills can be demonstrated only in a meaningful context. What is true of learning is also true of assessment for learning. If we test children individually on things that are actually done with others (e.g., using language) or are best used in real-life situations, then we should create assessment tasks that take these factors into account.

The emphasis on competencies—usually defined as skills, knowledge, and attitudes in a particular domain or field of learning—has tended to treat them as things one acquires and possesses, rather than as ongoing processes that manifest only in certain situations and under certain conditions. The idea that learning can be transferred from one domain to another is an oversimplification. In reality we always have to learn to apply skills in new situations. What we do transfer are dispositions to learning, habits of thinking and perception, and sensitivity to parameters. Capabilities are complex sets of dispositions and skills that enable us to perform complex tasks such as working with texts in foreign languages, carrying out an experiment in chemistry, or performing a play.

What we do transfer are dispositions to learning, habits of thinking and perception, and sensitivity to parameters.

We therefore need to set students tasks that enable them to use and develop a range of dispositions and skills in different authentic, complex situations—i.e., we need to have flexible forms of assessment. Here are some general suggestions in this direction:

- The task has to be clear to the students.
- The task must afford opportunities to use the competence appropriately.
- The task should involve some kind of performance—that is, a real task that is located in a meaningful context and that accomplishes something that the students experience as relevant. In Waldorf terms, this means that the task must engage the pupil's will.
- The assessment should occur over time, preferably across multiple contexts in which each context affords the use of the desired competence in different ways.
- Both the learner and the assessing teacher need to be clear about what kinds of evidence are required. Thus the assessment should be based on clearly formulated criteria.
- The process of assessment is also formative and should empower the learners by helping them to understand their learning better.
- Clear responses and suggestions, based on the evidence, make the achievements and the next steps clear to the students.
- When making an overall judgment about competency, several episodes over a period of time need to be included and several assessments taken into account.

Methods of monitoring and assessment

Some monitoring is informal, simply residing in the teacher's consciousness. However, the more students one teaches and the more classes one has, the harder this gets and the easier it is to form unconscious judgments of students that may not be correct, that might privilege some and discriminate against others. Few teachers admit to this, but I can say from personal experience, backed up by research, that this happens. (Kelly, 2011)

Therefore, some kind of notation and record-keeping is essential. This usually takes the form of a class list in which a number of aspects are simply recorded, such as attendance, punctuality, active and constructive participation in lessons, homework, completion of assignments, outcomes

of formal summative assessments, ongoing issues, learning support reports, outcomes of pupil case studies, and other relevant information.

The time-scale for this information depends on what is important and useful. There is little point in regularly documenting things we either don't need, can't understand, or are unable do anything about (e.g., ongoing medical conditions). This information makes sense only if it contributes to our understanding of the student.

School reports: At the end of each year, reports are given. In my school they come in the form of texts (not grades) that characterize the student and highlight her strengths and weaknesses. There is an evaluation of the student's participation in school, including attendance and punctuality. The student's social behavior is described. There is a short summary of the curriculum for each subject and an outline of the main learning outcomes. The student's attainments are noted and a formative comment about how she worked, what she found easy and hard is added. Finally, the next steps in overall learning, attitudes, and behavior are individually formulated for each student.

The report is aimed at informing parents and external parties rather than the students (who should already know this information directly and the report should contain no surprises), though for younger classes it can also be addressed at least in part to the child personally. It seeks to be fair and objective whilst demonstrating interest in the student. Some teachers err on the side of detailed perfection, either submitting their reports late or causing themselves stress in order to finish them on time. Others err on the other side, using standard formulations and not really addressing each individual. If good records are kept, writing a report should be straightforward

because it is a gathering of evidence and judgments made over time.

In high school, students should have the opportunity to read through their report and respond to it, taking up with their teacher any points they don't understand or agree with. If examinations or other external assessments (e.g., testimonials from work experience or projects) have been taken, these outcomes should be added to the report.

As research by Waldorf teacher Heidi Engel has shown, written reports enjoy several advantages over grades and competence profiles. (Engel 2013) A written report can be sensitively phrased to take account of the individual student. This is particularly relevant for weaker students, especially in cases when remedial measures are recommended. These can be framed in positive supportive terms that also value the efforts made by the learner (even though the outcome is normatively weak). Bald or generic comments lacking any context— such as "More practice needed" or "Must try harder"—are generally unhelpful and may reinforce an already poor sense of self-efficacy in the student. Just as unhelpful are unqualified cliché comments, such as "Good." A comment like this says nothing to those who made significant progress and further downgrades those who did not. Heidi Engel pleads for an assessment culture that is sensitive to the personal relationships (the how and the who), since these are so important in learning and development.

> **A written report can be sensitively phrased to take account of the individual. This is particularly relevant for weaker students.**

Competence portfolios: Increasingly, Waldorf schools are using portfolios to record the competencies young people develop through informal learning situations, such as work-experience practicums, projects such as the twelfth grade project, class plays, and other projects. The *European Portfolio Handbook* explains how these portfolios can

be structured. (Koch, 2010) As well as serving as self-assessments, competence portfolio documents can also contain external reports on young peoples' attainments. They are often used to support university applications, training opportunities, and work.

Naturally occurring evidence: This is evidence derived from activities in a range of real contexts and in everyday life over a period of time. This means that evidence should not be derived from contrived assessment events or artificial classroom situations. Evidence can be collected in any form that relates to the competence or capability in question. This includes work that students produce during the normal course of a lesson. The examples can be copied and collated over time. This is easiest to gather with written work, or with products that can be photographed, recorded, or filmed, but the same principle applies to oral performance (a presentation or participation in discussion, responding to or giving suggestions, asking questions and sustaining dialogue). The important thing here is that the teacher documents the occurrence, using criteria. Having colleagues collaborate and mutually confirm these judgments is important.

"Naturally occurring" does not mean "spontaneously occurring." Teachers need to design meaningful tasks that create opportunities for students to demonstrate their competence without creating artificial or contrived situations. And they must be able and willing to document these occurrences. This approach is similar to what Ciborski and Ireland report as performance assessments of experiential learning or assessment while learning is taking place, though actually what they describe is assessment when teaching or instruction is taking place. (Ciborski & Ireland, 2015, 85–87) Evidence of learning is not just how children behave during teaching

> **Teachers need to design meaningful tasks that create opportunities for students to demonstrate their competence, without creating artificial or contrived situations.**

episodes, such as participation. Again, it depends on what we understand as learning.

Using grades (i.e., rating using a scale of achievement): How we offer reflections and suggestions on the basis of assessment is crucial. Giving a student grades at any age is meaningful only when they are accompanied by reflections that show the student the way forward. Grades without commentary indicate the level of attainment at the end of a learning process—for example, when a student leaves school. Even then, they are better accompanied by some formative comment highlighting strengths, weaknesses, and concrete suggestions as to how the student can improve her performance.

Waldorf schools are famous in some quarters (for instance, widely in Europe and occasionally in North America) for not giving pupils grades, even in the upper years of their education. Unfortunately, this reputation is not always deserved. Many teachers give grades from the upper elementary school onward (sometimes only in the form of the results of tests, and the pupils treat them as grades). There are many problems with grades, and it is a very well-researched area of education. It is well known that different teachers will grade the same work differently, ranging from 1 to 4 points on a scale of 5, even when using the same criteria (though often these are implicit and not explicit). Texts are harder to grade than math exercises, but even in math there are surprising variations. Teachers mark differently if they know the students, compared to assessing them "blind." Sympathies and preferences, often unrecognized, play a role in how grades are awarded.

The notion that grades motivate and tell students where they stand is not borne out by research. The student who always gets top grades receives little incentive to expand knowledge

or skills, the student who always gets middle grade begins to feel very ordinary, and those who regularly get low grades already know they are weak (perhaps even "bad") students.

Grades offer little incentive and have a negligible effect on learning. If you give students grades and written reflections, they look only at the grade. Hand back work you have spent hours correcting and annotated with a grade and, in all likelihood, the student will look only at the end grade, perhaps comparing it with those given to classmates. In consequence, the student has learned nothing from the work the teacher has invested in assessing the assignment. All the evidence points to the conclusion that a teacher's response in the form of commentary is far more effective than any grades. (Hauschild 2014) It is very simple: Grades are not what people think they are.

The notion that grades motivate and tell students where they stand is not borne out by research.

Conclusions

Good practice occurs when self-reflective teachers assess their own teaching in terms of its effect on student learning and development and then share their evaluation with colleagues on a regular basis. In this way individual insights are tested collegially and flow back into the classroom. Good assessment can lead to transformation in both teaching and learning. When teachers use assessment intelligently to support learning and share it with their colleagues, one can speak of the school as a learning organization. If this is done effectively and transparently, the school can be considered accountable.

References

Ball, S.J. (2012). *Global Education inc. New policy networks and the neo-liberal imaginary*. London and New York: Routledge.

Ciborski, S. & Ireland, H. (2015). *Assessment for learning in Waldorf classrooms: How Waldorf teachers measure student progress toward lifelong learning goals.* Academica Press: Palo Alto, CA.

Drummond, M.-J. (2008). "Assessment and values: A close and necessary relationship," in S. Swaffield (Ed.), *Unlocking assessment: Understanding for reflection and application*. London and New York: Routledge, pp. 20–35.

Engel, H. (2013). *Verbalzeugnisse als Teil des schulischen Feedbackverfahrens beziehungssensibel gestalten*. (Masters), University of Osnabruck.

Hattie, J.A.C. (2012). *Visible learning for Teachers: Maximizing impact on learning*. London and New York: Routledge.

Hauschild, J. (2014). "Sind Schulnoten noch zeitgemäß?" *Spektrum der Wissenschaft (Psychologie/Hirnforschung)*, 09.10.2014. doi: http://www.spektrum.de/news/sind-schulnoten-noch-zeitgemaess/1312119.

Hipkins, R. (2007). "Assessing key competencies." In M. o. Education (Ed.). Wellington, NZ: Ministry of Education.

Keller, U.L. (2012). *Gerechte noten gibt es nicht (Just grades don't exist)*. Sinzheim: Via Interna Verlag.

Kelly, P. (2011). "Unconsidered activity, craft expertise and reflective practice in teaching," in *Reflexive practice: International and multidisciplinary perspectives*, 12(4), pp. 557–568.

Koch, T. (2010). *European portfolio certificate handbook*. Brussels: European Council for Waldorf Education/Comenius Lifelong Learning Programme.

Largo, R.H. (2011). *Kinderjahre: Die Individualität des Kindes als erzieherische Herausforderung (Childhood: the individuality of the child as educational challenge)*. Munich: Piper Verlag.

Mepham, T. & Rawson, M. (1997). *Steiner Waldorf education in the UK: Aims, methods and curriculum*. Forest Row, UK: Steiner Waldorf Schools Fellowship in the UK and Ireland.

Rawson, M. (2005). "The role of evaluation and examinations within Waldorf education within different age groups." In D. Mitchell (Ed.), *Evaluation, homework and teacher support*. Fair Oaks, CA: AWSNA Publications, pp. 25–36.

_____. (2010). *Educational quality development*. Eindhoven, NL: Mercurius International.

_____. (2011). "Democratic leadership in Waldorf schools." *Research in Waldorf education*, 2(2), pp. 1–15.

Rawson, M. & Richter, T. (2000). *The educational aims and content of the Steiner Waldorf curriculum*. Forest Row, UK: Steiner Schools Fellowship Publications.

Rumpf, H. (2010). *Was hätte Einstein gedacht, wenn er nicht Geige gespielt hätte? Gegen die Verkürzung des etablierten Lernbegriffs*. Weinheim and Munich: Juventa Verlag.

Scharmer, O. (2009). *Theory U: Leading from the future as it emerges – The social theory of presencing*. San Francisco: Berrett-Koehler.

Schieren, J. (2012). "The concept of learning in Waldorf education." *Research on Steiner education*, 3(1), pp. 63–74.

Swaffield, S. (2008). *Unlocking assessment. Understanding for reflection and application*. London: Routledge.

Swaffield, S. & Dudley, P. (2011). *Assessment literacy for wise decisions* (3rd edition ed.). London: ATL the education union.

Thomas, R. (2005). "Evaluating, judging, testing and learning." In D. Mitchell (Ed.), *Evaluation, homework and teacher support*. Fair Oaks, CA: AWSNA Publications, pp. 19–24.

Verhaeghe, P. (2015). *What about me? The struggle for identity in a market-based society*. Melbourne and London: Scribe Publications.

Martyn Rawson *is a high school teacher at the Waldorf School in Elmshorn, Germany, and also in the Masters Program at the Waldorf Seminar in Kiel. He is a guest professor at the Freie Hochschule Stuttgart and London Waldorf Teacher Training Seminar, as well as being involved with teacher education in India. A research associate at the University of Plymouth Institute of Education in the UK, he has published widely on Waldorf themes, including co-editing* The Tasks and Content of the Steiner-Waldorf Curriculum *(Avison & Rawson, 2014).*