Spring 5-18-2023

Alternative Assessment in Education

Eugene Jack Constandaki
University of New Hampshire, Durham

Alternative Assessment in Education

Eugene Constandaki

Introduction

Educational progress has been oversimplified into a letter grade which students receive at the end of their courses. This can cause students to believe that the only thing that is important when learning is their grade: their mentality is "learn for the grade" when it really should be "learn to become interested and educated." It's easy to get lost in this mindset, in fact, I struggled with this in my years as a student. Learning was so stressful and tedious that it made me lose sight of the end goal: finish school imbued with all the knowledge and practical skills I had been taught. It was difficult to be interested in the subjects I was learning about because I knew my education was ultimately boiled down to a letter, which was all I needed to move on to the next step. Educators should be enthusiastic about inspiring interest, but the grading-heavy curriculum currently in place creates a large obstacle. By lessening the concern students feel due to their grades, we can make true learning more achievable for them. Obviously, testing of some sort must still occur, otherwise how would we know what information is being retained by students? Thus, my research question is as follows: how can we assess student's knowledge while easing the concern they feel with grades? I will put the knowledge I have gained from this research into practice in my own teaching, as I plan to become a high school physics teacher after graduation.

Competencies as a Form of Assessment

Spaulding High School, located in Rochester, New Hampshire, has adopted a competency-based grading system. I interviewed principal Justin Roy, administrative assistant Angela Green,

and physical science teacher Caitlin Gately, who were eager to share their insight on the system that they have had in place since 2010. They were enthusiastic about the change from a traditional grading system, noting it as a step in the right direction that offers students multiple opportunities to show what they know and what they are capable of in an ever-changing system. Mr. Roy believes that they are "on the leading edge in New Hampshire along with a few other schools" due to their use of competencies and other strategies which will be discussed.

Each course at Spaulding High School has a set of five or six competencies which are clearly outlined and made available to the students. Each of the competencies is relevant to a unit that is taught in the course and is split up into smaller parts to create a more logical evaluation criterion. Each competency is evaluated with a familiar scale: A (advanced), B (beyond competent), C (competent), and not yet competent. When being evaluated, the teacher scans the competency rubric and finds which description best fits the student. For example, when gauging a student's competency in "Newton's Laws of Motion with a concentration on the Second Law," the teacher has four options: "[the student] can apply Newton's Second Law to novel situations and predict the motion of objects using complex mathematics" (A), "[the student] can appropriately apply Newton's Second Law to various scenarios and support it mathematically" (B), "[the student] can summarize Newton's Laws specifically the Second Law. [The student] can support my summary by citing specific details and examples along with basic mathematical reasoning" (C), and "[the student] can state Newton's Laws conceptually" (not yet competent). The system is quite different from the traditional grading system that was in place before, and required some adjustment from faculty, students, and parents.

Change is never easy, and when asked how the students and faculty responded to the transfer to the competency structure, Ms. Gateley explained that the competency system has

boasted huge advantages for the average student. Having a rubric allows students to view a map of what they needed to do in the past and how to improve on those subjects to get to where they need to be. The more studious few are typically locked into the idea of striving for high marks and doing well with respect to their peers, which presents them with a small obstacle as competencies dampen the competitive nature of the traditional grading system. Surprisingly, in a developing system, parents were perceived to have had the most difficulty with understanding competencies, as they grew up in a system that only offered a grade in their learning process and want to know how their student compares to the others in class, which is difficult to see when competencies are being used. "Initially there was a lot of pushback; I don't think that a lot of them understand it even now, but they have grown up with it, [as] the kids in high school have had it as well as their parents all through middle school and for most of elementary [school], but initially a parent wonders how their child ranks among their classmates," said Mr. Roy. With competencies, students aren't assigned a grade, which eliminates that type of competition. Instead, they are placed into the section of each competency that best matches their work.

Mr. Roy offered that it is impossible to find a system that everyone loves, but that is a side effect of never being able to appease everyone. Ms. Green brought up how "teachers who have moved on to different districts who have used the competency system have given feedback that this method is true to learning and that is our ultimate goal... I love the fact that we're offering kids multiple opportunities to show what they know and are able to do." One of these opportunities that the students at Spaulding high school have is 'move-on-when-ready.' For most courses, students may choose to either proceed with the material at a standard pace or a bit faster. This mainly benefits those students who like to get ahead and encourages those who don't to have an adaptable pace. This allows them the freedom to build collaboration, communication, and critical-thinking

skills, while placing more responsibility on the students to regulate their discipline in meeting deadlines, which as Ms. Green says, "prepares students for the workplace now and in the future."

It seemed to me that having students moving at different paces while also having to keep unique competencies in mind when grading would be very difficult for teachers, and that students may struggle with keeping track of what competencies they are expected to fulfil, especially because the competencies for, say, Physics, are different from the ones for Mathematics. Ms. Gately added that "reading the rubric and comparing their input from the rubric is difficult for the kids. Having clear expectations on the project and clear rubrics are the most important things to help the kids with that." If there are parts of the competencies that are difficult to understand, it would be helpful for teachers to go over them in class or even create an alternative rubric that presents the criteria from the competencies in a less–technical way which would make understanding the expectations for learning in a unit easier for students and teachers alike. I also questioned whether some students would take advantage of 'move-on-when-ready.' While a valid concern, students are not permitted to finish a course late. 'Move-on-when-ready' is intended to be an incentive for students to take initiative and finish their work ahead of time, thus establishing time management, planning, and foresight skills in students who wish to progress at a pace faster than the standard of the class. Ms. Gately has noticed that the hardest part for students is "having the self-discipline to meet deadlines and be able to be self-directed," while communication and collaboration comes more naturally to them. Those students who don't take advantage of 'move-on-when-ready' are pushed by teachers on a more individual basis to work on reinforcing their understanding of concepts and establish a more concrete accountability structure. "Teachers are now giving feedback on how self-directed students are and how they are at goal setting and meeting

their goals," said Ms. Green. While there are challenges to be faced when morphing into a new system, competencies offer some great advantages.

One of the biggest advantages of competencies is that they allow students to clearly see what their strengths and weaknesses are when looking back at their evaluation and use that information to fortify their learning process, which is not something they were able to do very easily with a letter grade. Competencies give students a new way to set goals, and when viewing their placement in past and present competencies, students can accurately measure their progress through the year. Spaulding High School allows students to 'relearn and reassess' their knowledge. 'Relearn and reassess' is the antithesis of 'move-on-when-ready:' students who learn at a faster pace can benefit from being able to finish material early while students who learn more slowly can relearn and reassess if they fail to meet the criteria during a certain period. If a student did not meet the minimum competency in one or more areas, they can look back, see exactly where they need to improve, study the material again and put their knowledge into practice by reassessing. This is made possible through competencies; if all they were given was a letter grade, it would be difficult for them to pinpoint exactly where they needed to improve. Teachers can look back at the placement of students in their past competencies, which gives teachers insight into where individual students might need extra help. No system is ever perfect, but I must agree that competencies are a fantastic step in the right direction.

Effect of Feedback on Performance and Interest

I am far from the first person to question the implications of grades or whether there is a better form of feedback to give students. Ruth Butler, professor of Educational Psychology at Hebrew University, documented a study titled "Enhancing and Undermining Intrinsic Motivation:

The Effects of Task-Involving and Ego-Involving evaluation on Interest and Performance" on the hypothesis that the fundamental motivation of students would be differentially affected by task-involving and ego-involving evaluation, and that both kinds of evaluation would promote both types of involvement. Task-involving evaluation was in the form of comments on the student's performance, while ego-involving evaluation was in the form of a grade. Understanding how the feedback students receive from evaluation affects their performance and intrinsic motivation to continue to be interested in a topic is critical in optimizing evaluation for education.

The pupils for the study were made up of, according to their most recent report card, the top and bottom twenty-five per cent of "twelve classes of fifth and sixth grade students, [who] were randomly assigned to one of three feedback conditions" (Butler, 1) after performing tasks which they had a high initial interest in: grades, comments consisting of one sentence which was related to performance only, or both. Students from this age group were chosen because, according to research done by John G. Nicholls, Doctor of Education and Psychology, "by fourth grade most pupils have achieved a differentiated concept of ability as a stable trait best assessed by comparison with others. At this point grades should induce a shift from task- to ego-involvement" (Butler, 3). In short, by the time students are in fourth grade, their mode of self-assessment has shifted to comparing themselves and their grades to that of their peers. This is important because it provides the researchers with a more stable source of data that complies with the ego-involving focus of the study. The study was broken up into three sessions: in session 1, pupils in all three groups were given two identical Tasks, A and B, which required the students to respond to questions in writing. After 10 minutes of working on Task A, pupils were asked to stop and begin Task B. After both tasks were complete, pupils were given an interest questionnaire to gauge their interest in the topics of the Tasks. Two days later, students were given their feedback in the form of a grade, comment,

or both. Pupils then took part in Sessions 2 and 3, which were spaced 3 hours apart. In Session 2, they were given new material and were asked to carry out Tasks A and B again and were given the interest questionnaire upon completion. In Session 3, after receiving their evaluation, they were given the third set of material and carried out Tasks A, B and the interest questionnaire once more, but this time they were told that their responses would not be evaluated. After the study, all participants engaged in a discussion about the experiment. Table 1, below represents the timeline of the study:

Day 1	Day 2	Day 3
Session 1		Receipt of Feedback
Task A (10 Minutes)		Session 2
Task B		Task A (10 Minutes)
Interest Questionnaire	BREAK	Task B
		Interest Questionnaire
		Receipt of Feedback
		3 Hour Mark
BREAK		Session 3
		Task A (10 Minutes)
		Task B
		Interest Questionnaire
		Discussion

Table 1

Finding what affects intrinsic motivation was crucial in this study, as the purpose of the study was to increase the effectiveness of education through interest for students. Butler states that "recent research on intrinsic motivation has consistently found that rewards undermine subsequent interest for initially attractive tasks, apparently by promoting attributions of task engagement to the reward rather than to pleasure in the activity itself" (Butler, 1). So, theoretically, by evaluating students with a grade, the students would take less pleasure in completing the task as they are crediting their performance to a grade (reward), and the task for which they are being evaluated

would lose its value. The use of grades would then involve the ego of the students, as they would focus their attention on demonstrating high ability or masking low ability with their marks. When they receive their scores, they would feel a "self-esteem-based pressure to achieve positive and avoid negative outcomes...While grades may maintain initial levels of interest and performance in high achievers who expect these to be self-enhancing, these should be undermined in low achievers" (Butler, 2). This is summarized by the experimental hypotheses: "(a) Post-test interest and performance on both tasks will be highest after receipt of comments at both levels of school achievement. (b) High achievers will score similarly on Session 2 interest and convergent thinking in all groups, while low achievers will score highest after comments; both high and low achievers will score highest on immediate divergent thinking after comments. (c) Subjects who received comments alone will recall these better than subjects who also receive a grade; changes in performance from pre-test to post-test will be related to the content of the comments received earlier in the Comments but not in the Grades + Comments condition. (d) Patterns of interest and performance on Sessions 2 and 3 will be similar in the Grades and Grades + Comments conditions" (Butler, 3). This hypothesis states that (a) both high and low achieving participant's scores and interest in the subject will improve after the receipt of comments, (b) the scores received by high achievers will be similar across the board in Session 2 while low achievers would benefit most from comments, (c) comments and their quality would proportionately affect the performance of subjects, and (d) the interest and performance of subjects in groups that receive grades in Sessions 2 and 3 will share a pattern.

A two-way analysis of variance (ANOVA) was done to assess the validity of data collected. ANOVA is used to determine whether the null hypothesis can be accepted, how likely variance between groups is, and the ratio of variance between groups to variance within groups. 'Two-way

analysis' means that the data analyzed has two variables: in this study, those variables were the feedback condition received and school achievement of subjects. ANOVA yields two values that are significant to variance: F-value and P-value. The F-value describes the ratio of variance between groups to variance within groups and the P-value shows how likely variance between groups is. For example, when analyzing the variance between the scores of high and low achievers completing Task B in session 1, ANOVA yielded $F(1, 125) = 6.01$, $P < 0.05$. This showed that, assuming all groups will produce similar results, the probability that the data collected from them differs between groups was less than 5% according to the P-value, while the ratio of variance between groups to variance within groups was 6.01:1. In different terms, the P-value tells us that variance between groups is unlikely, while the F-value tells us that there was a massive ratio of variance between groups when compared to variance within groups. Clearly, this analysis does not match and so, the null hypothesis must be rejected, meaning, when analyzing variance between the scores of high and low achievers completing Task B in session 1, it cannot be assumed that high and low achievers will produce similar results and the data collected from that portion is statistically significant. All analyses of variance which resulted in a P-value higher than 0.05 were noted as having 'no significant main or interaction effect' and the values for them were not reported.

The relevant resulting data from the study is displayed below in tables 2 and 3. Table 2, which was imported directly from Butler's study, shows the means and standard deviations for performance scores of high and low achievers in all groups of feedback condition for each session and task. Scoring is based on the categories defined by Torrance and Templeton, who designed a template for testing different forms of thought. In this scoring system, the higher the score, the better the performance of the group. The conclusions from this data are comparative, so the

individual values are not as important as the difference between values. I created table 3 which describes the ratio of difference in performance and perceived interest between groups to difference within groups using the significant ANOVA results that were reported in the study. By interpreting data from both tables 2 and 3, the following conclusions can be made: the ratio of variance of scores between groups with different levels of school achievement to variance of scores within those groups were 13.02 and 6.01 for session 1 tasks A and B, 14.15 for session 2 task B, and 57.48 and 75.73 for session 3 tasks A and B, respectively. The comparison for the high achiever group in session 2 task A was not deemed significant, meaning the ratio of variance of scores between the high achiever and low achiever groups and variance of scores within the high achiever group was close to 1. This first observation indicates that there was a constant significant difference between the scores of groups with different levels of school achievement and scores within those groups. Another observation that can be made is that the ratio of variance of scores between groups with different feedback conditions to variance of scores within those groups was 29.6 and 44.36 for session 3 tasks A and B respectively. This, coupled with information from table 2, shows that by the end of the test, students in the comments only group performed significantly better than those in other groups. These ratios already seem very high, which shows how stunning the ratio of perceived interest between the comments group and other feedback condition groups was to variance within those groups in session 3 for tasks A and B: 245.4. Clearly, receiving only comments not only sustained the interest of the students involved in the study, but increased it tremendously compared to the interest of students in the grades or grades plus comments groups.

		Comments group		Grades group		Grades plus Comments group	
		High	Low	High	Low	High	Low
Task A							
Session 1	M	18·77	10·14	19·64	9·86	19·60	9·64
	SD	6·63	5·02	8·67	4·88	5·80	5·64
Session 2	M	25·36	17·86	24·95	12·50	16·77	6·55
	SD	9·33	8·70	9·91	7·86	5·61	4·73
Session 3	M	24·27	13·50	16·45	8·59	11·82	5·82
	SD	6·70	7·79	9·34	6·41	5·26	5·40
Task B							
Session 1	M	19·36	10·27	19·68	10·32	18·68	10·12
	SD	5·35	4·01	7·36	3·71	6·60	3·63
Session 2	M	25·59	17·64	16·09	12·59	14·36	11·95
	SD	9·39	10·16	6·63	5·94	5·17	3·71
Session 3	M	24·95	14·27	14·91	8·50	14·95	9·18
	SD	5·38	4·61	6·63	3·07	5·83	3·40

Table 2

Session	Task	Analysis	F-value	P-value
1	A	Performance vs School Achievement	13.02	0.001
1	B	Performance vs School Achievement	6.01	0.05
2	A	Performance of Low Achievers vs Feedback Condition	14.54	0.001
2	A	Performance of Low Achievers Receiving Grades vs Other Feedback Conditions	5.84	0.05
2	A	Performance of High Achievers Receiving Grades vs Other Feedback Conditions	10.85	0.001
2	B	Performance of High and Low Achievers vs Feedback Condition	17.67	0.001
2	B	Divergent Thinking of High and Low Achievers vs Feedback Condition	33.96	0.001
2	B	Performance vs School Achievement	14.15	0.001
3	A	Performance vs Feedback Condition	29.6	0.001
3	A	Performance vs School Achievement	57.48	0.001
3	A	Performance of High and Low Achievers Receiving Comments vs Other Feedback Conditions	98.68	0.001
3	A	Performance of High Achievers Receiving Grades vs Those Receiving Grades Plus Comments	11.8	0.001
3	B	Performance vs Feedback Condition	44.36	0.001
3	B	Performance vs School Achievement	75.73	0.001
3	B	Divergent Thinking of High and Low Achievers vs Feedback Condition	160.1	0.001
2	A and B	Perceived Interest of High and Low Achievers vs Feedback Condition	27.44	0.001
1 and 2	A and B	Perceived Interest of High and Low Achievers vs Covariate	37.21	0.001
2	A and B	Perceived Interest of Grades and Grades Plus Comments Groups vs School Achievement	13.56	0.001
3	A and B	Perceived Interest vs Feedback Condition	49.54	0.001
3	A and B	Perceived Interest of High and Low Achievers of Other Feedback Condition Groups vs Comments Group	147.6	0.001
3	A and B	Perceived Interest of High and Low Achievers vs Desire for Extra Tasks	66.66	0.001
3	A and B	Perceived Interest of High and Low Achievers in Comments Group vs Other Feedback Condition Groups	245.4	0.001

Table 3

The final results showed that over the course of the study, the performance of both high

and low achieving students in the comments group improved for both tasks, with their highest

11

mean scores being in session 2 and their session 3 mean scores being significantly higher than those of session 1. The performance of high and low achieving students in the grades group improved from session 1 to session 2 in task A, but due to lost interest in the topic as a result of crediting their performance to a grade, their mean scores were at their lowest after session 3. The same result occurred in task B, with the exception that only low achieving students improved from session 1 to session 2; the high achieving students' mean scores declined throughout. The mean scores of both levels of achievement in the grades and comments group declined throughout in task A, while in task B, the high achieving students performed best in session 1, and worst in session 2, and the low achieving students performed slightly higher in session 2 than in session 1 and lowest in session 3 on average. If, in this study with these students, performance and interest were at their highest after the receipt of only comments as a form of feedback, then comments should at the very least be administered to students more frequently in a system where grades are predominant. As I mentioned in the last section, the use of competencies is becoming more common in schools, but why stop at competencies? Though more tedious, connecting with students on a more individual basis through comments on classwork would be greatly beneficial to the growth of students. Competencies offer a baseline for students to recognize their strengths, weaknesses, progress, and goals, and are convenient for managing feedback in large classes. However, recognizing a student's progress individually and explaining to them how they can learn from their mistakes in more detail would be a step further.

Waldorf Education

There exists a method of education in which students are taught through experiences, assessed through trials, learn from the mistakes they make and feedback they receive from

assessment, and grow as a person as well as a curious student. This method of education is called Waldorf Education. Students at Waldorf schools are heavily encouraged to make mistakes on their path of learning and use their trials and assessments as learning opportunities. Where traditional schools offer a competitive and ambitious environment, Waldorf schools foster an environment of cooperation and interest.

Martyn Rawson, class teacher in York, United Kingdom who also taught English, art history, and anthropology in Germany, wrote a journal on "The Role of Evaluation and Examinations within Waldorf Education within Different Age Groups." In this journal, he describes the environment in which Waldorf ideologies are used and why they work. To understand how assessment is used as a tool in Waldorf schools, we must understand the natural development of a student's mind, the role of teachers, and recognize the differences between Waldorf evaluations and the assessments that most traditional students are used to.

All humans, and therefore all students, are naturally curious. From a young age, children attempt to understand the world better through their experiences, successes and failures. They constantly generate questions that only knowledge can satiate, and as they gather knowledge, they eventually reach a point where they feel a need to be challenged and tested. However, that need "is not met through forms. Rather they seek their trials through all manner of risk-taking activities" (Rawson, 25). In a school setting, students should feel challenged, but if not through forms and assessments, how can this be accomplished? Waldorf students are encouraged to engage in self-evaluation, where "pupils can reflect on their learning, on their strengths and weaknesses, what they enjoy, [and] what goals they set themselves. The pupils can focus on schoolwork, behavior, attitudes and their own competencies" (Rawson, 31). Goal setting is an essential part of the learning process, as it gives students tangible heights to reach as well as achievements to celebrate.

Rawson answers concerns regarding the capabilities of young students by describing pupils as "quite capable of identifying their own relevant criteria and goals and assessing their own achievements" (Rawson, 34). In fact, Rawson even implies that pupils begin engaging in pedagogical thought starting in fifth and sixth grade. "It is in the fifth grade that one can first begin to talk openly with the pupils about the learning process itself. One can discuss what helps and what hinders the learning process. From this age onwards the children can begin to reflect on their own feelings and their own learning process...They begin to be interested in how different people are and how people learn differently, things they have intuitively known before but can now begin to observe and verbalize. In sixth grade the question of learning gains a new dimension with the awakening of new faculties that make causal thinking possible" (Rawson, 33). Perhaps Waldorf schools provide some sort of environmental catalyst which amplifies the ability of students to question their mode of learning and ask themselves how they learn and what they have learned compared to what they were taught. I would argue that this is true, and the catalyst provided is due to using evaluation as a tool to strengthen the learning process. I will discuss this further when comparing evaluation in Waldorf schools to that in traditional schools.

As students fail in the process of eventually succeeding in their goals, they reach the most influential part of their learning journey, which is making and realizing their mistakes. Up to this point, the purpose of the instructor has been to educate, set goals with, and challenge students. It is, obviously, much more involved than that. In Waldorf schools, teachers evaluate students with a unique but productive purpose: "to strengthen the learning process through evaluation" (Rawson, 27). In this sense, evaluation is defined as assessing the value of, comparing, and making visible the knowledge of students while appreciating its unique qualities. Evaluation is not made in the form of an exam, but rather in the process of monitoring a student's progress and recording their

achievements towards levels of competence. These evaluations come in many different forms; if learning can be achieved through different forms of instruction (such as visual, auditory, reading/writing and kinesthetic (VARK)), so too can different forms of assessment accurately evaluate students. Rawson provides an example of this in Table 4 below. Monitoring "require(s) that we have developmental milestones or thresholds in mind, expectations that tell us that children usually reach this stage of development around this time or in certain sequence" (Rawson, 30). This is very different from testing, which "requires that we construct a specific situation in order to test whether certain skills have been learned, e.g., through a vocabulary test or a written assignment under certain conditions. If monitoring is passive observation of that which unfolds, testing is active" (Rawson, 30). While both are important, monitoring is usually left behind in the traditional system.

Different learning forms need different forms of evaluation

Learning mode	Characterization of the learning mode	Possible form of evaluation
1. Imitation	Imitation presumes the presence of competent role models. Learning through participation, copying, repeating later what one has witnessed or been deliberately shown, some imitation can be explicit. Preconditions include interest, awareness, certain level of relaxation and sense of well-being (sound emotional and physiological basis needed).	Documented observation e.g., in profiles
2. Learning through explicit teaching and instruction	Usually through planned deliberate showing, demonstrating, describing or exemplifying of the stages and steps of specific processes. Best done in context. Preconditions include willingness and ability to concenrate for a given time span and sound visual, auditory, motor abilities (relaxed awareness). It is necessary that the learner is stimulated in his feelings and can identify with the content and learning process (i.e. he feels that it is relevant and important).	Through recall, specific questioning of detail by the teacher of the learner, through application in recall and and through application in problem solving
3. Repetition	Learning through repeating of activities, processes (ideas and words) what has been experienced. Needs variation and focus to maintain interest and momentum. Preconditions: rhythm, routine, artistic structure or content, relevance and ability to apply what has been learned to specific contexts.	Activity can be witnessed and documented.
4. Learning through storytelling	The teacher (or pupil) presents complex content in narrative form verbally. Preconditions: listening, empathizing skills (both storyteller and listener)	In verbal or written recall, in creativity ` by the storytelling
5. Learning through deliberate observations, including focused and focusing questions	Observational tasks are given either with or without specific focusing in a specific area of experience (e.g., text, series of problems, specific setting, etc.). Preconditions: meaningful context, interest and identification (e.g., also relevance), good observational skills	Set tasks and assignments with clear instructions, tools (both techniques and technology), written or verbal (usually with a spectrum of possible correct answers)

6. Learning through discovery	Outcome: open exploration of phenomena, no predetermined outcome. Emphasis and reflection of experiences. Also learning through doing in practical contexts, where the outcome is specific (a task that needs doing) but the experience is also open. It is also open what forms are used by the learners to give expression to their discoveries. Preconditions: confidence, good basic skills (conceptual, literate and oral, organizational), motivation	Self-selected method of reflection and reporting: diaries, portfolios, short talks, exhibitions, etc., competence proofing
7. Self-directed learning with support	The learner chooses the topic/theme, carries out the work alone (or in a team) and receives support on re-quest from the teacher. Precondi-tions: (as above) plus the ability to ask for and knowledge of how to ask (or seek)	Portfolio, presentations, dialogue, competency profiling
8. Independent working	The learner asks his own questions, sets his tasks, chooses his own methods and motivation Precondi-tions: all the above!	Portfolio, presentations, dialogue, competency with portfolio

Table 4

With all this in mind, how does testing in a Waldorf School differ from that in a traditional school? Waldorf schools use testing as evaluation but more importantly, as a tool to strengthen the learning process by focusing on mistakes made during evaluation. Learning does not end once evaluation is administered, rather, it is just at the beginning. Rawson brings up an excellent point about the underlying issue of exams: "in uncertain times, in which all forms of social security are being progressively withdrawn or undermined, the pressure from parents and pupils to quickly and efficiently gain the security of a passport to the next round of qualifications grows apace" (Rawson, 26). Students in the traditional system are constantly reminded that their passing grade is all they need to make it to the next level, which causes them to focus on their grade rather than their educational experience, which serves to undermine their intrinsic motivation to learn and interest

in the topics they are supposed to be learning about. Additionally, understanding that something as small as a grade can have such a large impact on a student's life as preventing them from progressing to the next grade causes unwarranted amounts of stress. This is true for state exams as well, which "generate such insecurity that they work to undermine the education right down into the elementary school, years before any heed need be paid to the actual requirements of the exams. Their psychological effect is much greater than their actual relevance" (Rawson, 26). In recent years, some countries have begun prescribing learning levels to pupils (much like competencies) as a form of accreditation, which takes the place of traditional exams. The lowest level would describe a pupil who is just being introduced to the subject, while the highest level would describe a pupil who is nearly completely competent in the subject and requires little advice or training from the instructor. Table 5, below, describes these learning levels:

Level	Loebell	Steiner, *Study of Man* lecture 9	Ability/ competence	State educational standards
A	Awareness/ noticing	Conclusion (willing)	Awareness, powers of perception	Reproduction
B	Commitment	Judgment (feeling)	Ability to apply what one has learned	Application
C	Evidential experience	Concept (thinking)	Ability to abstract	Transfer

Table 5

There is a way to evaluate students without generating insecurity or building a stressful environment that causes students to be more focused on a grade than on the material. Evaluation should allow students to demonstrate their competence and abilities and give them an opportunity to learn not only in the classroom but also in real life situations. All modes of evaluation in Waldorf

schools result in students receiving "feedback on their work and development" (Rawson, 26). One such mode of evaluation is through portfolios.

The main form of assessment in Waldorf schools is the portfolio which each student creates. This portfolio is a conglomerate of work the student has produced, their self-evaluations on those works, and recognitions of goals they have reached in their personal learning process. Student-created portfolios provide a way for pupils to comment on their own work. "This allows for a much more individual presentation of attainment in which the strengths and weaknesses of the individual in a varied and focused form are far more concrete in its quality of expression than with formal marking or grades. A portfolio tells us far more about the learning process and development of the individual than tests or other forms of traditional exams" (Rawson, 34). I strongly believe that having a portfolio is a much greater representation of educational progress than a grade.

Lesson Plan and Rubric

Below, I have created a lesson plan and rubric for a class on forces. As I plan to become a teacher, I will use this as a template which I will build on in the future. This template draws on concepts I have learned through working on this project, such as competencies, VARK, the effect feedback has on the performance and interest of a student, and methods used in Waldorf schools. It begins with a bullet-point list of topics for a two-week long lesson arranged in an order I find to be logical and true to an experiential learning process. General teacher goals are outlined next, with an explanation of the lessons following. The final piece is the rubric, which includes a note on how the rubric is set up and what assignments would be evaluated using it.

- Unit on Newton's Laws and forces including friction but not tension or springs

  - Lesson plan:

    - What is a force?

    - Types of forces:

      - Force due to gravity

      - Normal force

      - Applied force

      - Friction (kinetic and static)

    - Free body diagrams

    - Forces interacting with each other:

      - Newton's 1st Law:

        - An object at motion/rest will maintain its posture until acted on by an outside force

      - Newton's 2nd Law:

        - Force equals mass times acceleration

      - Newton's 3rd Law

        - For every action there is an equal and opposite reaction, third law pairs: if an object exerts a force on another object, then the other object also exerts an equal and opposite force on the original object

    - Units

  - Goals:

    - Connect with students and their experiences

- Present content in different ways: VARK (visual, auditory, reading, and kinesthetic)

  - Hands on/experiential learning to start, maybe with a guided question (physical examples, kinesthetic)

  - Lecture portion for visual examples with plenty of drawn pictures/FBD's (visual)

  - Split into groups to complete activity with monitoring (group learning)

  - Show equations and how they work (logical)

  - Set individual goals with students

- Lesson:

  - Start class by observing a block on a surface while students are watching you (visual). Try to move it. Observe what happens when you don't apply enough force to move it, when you apply enough force to move it, and when you apply so much force that the block continues moving after it was pushed until it slows to a stop. Split students into four or five groups to attempt the same observations (kinesthetic). After about 5 minutes, each group writes down a question based on their observations on provided individual white boards or in their notebooks (group learning): why does it stay still until it is pushed? Why does it move when pushed? Is there a minimum amount of force required to push it? Why doesn't it keep moving once pushed? Do other objects obey the rules of this block? Write these and other questions that arise down and discuss them (reading, auditory). When discussing a

specific question, write it down on the board and use a physical example to demonstrate the question. At this point, you have engaged questioning through experiential learning, using kinesthetic, visual, reading, and auditory cues. Introduce the concept of force and ask students what forces might be acting on the block. Introduce the concepts of pushing force, force due to gravity, normal force, kinetic and static friction, and their symbols. Assign homework describing what those forces are, how they interact, and what individual rules they follow.

▪ Introduce the tool of free body diagrams to help students visualize what is happening. Have the students make their own free body diagrams for a block that is at rest on a flat surface with no external forces other than that of gravity and the normal force, then one with a horizontal pushing force added, and then one after it is pushed and before it comes to rest. Explain that the free body diagram for the block before it is pushed is the same as the free body diagram for the block after it has been pushed and comes back to rest. Go through those example free body diagrams on the board to ensure everyone is on the same page. Address questions. Introduce Newton's laws of motion to help explain that what has been observed in the beginning of class can be replicated with other blocks on other surfaces. Write his three laws on the board and as you go through them, provide a physical example of a block experiencing the law which you are describing. Explain the units associated with force (N=kg*m/s^2). They should feel familiar to students as a result of the previous lesson. Assign homework for students to describe

Newton's laws of motion in their own words and draw out free body diagrams for the examples that were gone over in class.

- Short quiz with low weight on describing what the aforementioned forces are and how they depend on one another, one free body diagram, and Newton's laws of motion in their own words. Mistakes made on this quiz will be a building block in preparing for future assessment. Lab in which students recreate initial examples of the motion of the block with their new knowledge of the laws of motion. Lab report with heavy weight where they tie their observations into the content of the lesson in a report. Assignment for students to look back on their quiz and see what they could have improved on based on quiz feedback.

- Self-evaluation in which students describe what they know how to do and where they could improve.

- Rubric:
  - {Homework on forces, homework on Newton's laws, quiz on all, lab report, and self-evaluation} all included in their portfolio, portfolio will be graded according to following rubric: (standards are based on Spaulding High School's competencies (A, B, C, NYC), students will be placed into these categories based on their work and will receive individual comments that explain what parts of their work indicated their placement.

Subject	Advanced	Beyond Competent	Competent	Not Yet Competent
Forces	Can set up a physical example and fully describe how all forces are acting on the block and how forces influence the acceleration of a body, to include zero acceleration.	Can fluently explain interaction between all forces and the objects they are acting on by any means. Can fully explain (by any means) how forces influence the acceleration of a body, to include zero acceleration.	Can define, identify and differentiate all forces in lesson.	Does not meet competent.
Free Body Diagrams	Same as beyond competent but can relate a multitude of free body diagrams to one another and recreate scenarios described by free body diagrams.	Same as competent but can also describe acceleration of the body based on how forces interact with the bodies.	Can draw free body diagram for a block sliding to rest on a horizontal plane and correctly label forces on it with their proper directions and relative magnitudes.	Does not meet competent.
Newton's Laws	Can relate Newton's Laws of Motion to real-world examples.	Can accurately explain Newton's Laws of Motion in their own words.	Can identify Newton's Laws of Motion in order.	Does not meet competent.

- Next lesson will delve further into Newton's 2nd law with a more mathematical approach.

- Encourage mistakes and questioning

- Self-directed assessment and critique

- Testing that is used not only as a form of evaluation, but as a tool for the learning process as described by Rawson

Conclusion

In summary, students feel an unnecessary amount of pressure due to the disproportional significance which grades hold. The importance grades have been assigned in traditional education causes students to learn to receive a good grade, not to become educated or satiate their natural curiosity. A grade alone is not enough to describe the facets, progress, and ability of a student. The study done by Butler shows that students who received comments outperformed students who received grades or grades plus comments, and that the perceived interest of students who received comments was much higher than that of students who received other forms of feedback. Testing is a tool used in Waldorf schools to verify what has been learned, and another chance for students to make mistakes and learn from them in each lesson. Waldorf schools rely heavily on the idea that the best way for pupils to learn is from their mistakes through questioning themselves and their educators. This procedure "awaken[s] critical faculties and interest in the learning process" (Rawson, 33), which reminds me of the 'move-on-when-ready' system that Spaulding high school has in place and the portion of Butler's study involving interest (recall that perceived interest was highest after the receipt of comments only, "combined interest of high and low achievers was highest after comments, $F(1, 125) = 147.6, P < 0.001$" (Butler, 9). Portfolios are a fantastic way for students to lead their own educational experience, reflect, learn from their mistakes, and recognize their progress towards their goals. I believe these strategies are the best way to assess students' knowledge while easing the concern they feel with grades.

Sources

Spaulding HS competencies:
https://docs.google.com/document/d/13TsG4q6CGUPH5fhpEzxO1XJHhA1Wo-Ge0pO9AFtHRfU/edit

TORRANCE, A. E., and TEMPLETON, 0. E. (1963). Minnesota Tests for Creative Thinking. Minneapolis: University of Minnesota Press.
https://sclfind.libs.uga.edu/sclfind/view?docId=ead/ms2372.xml%3Bbrand=default

RAWSON, M. Waldorf: The Role of Evaluation and Examinations
https://www.waldorflibrary.org/images/stories/articles/WJP5_rawson.pdf

Geert van den Berg, R. ANOVA – Super Simple Introduction https://www.spss-tutorials.com/anova-what-is-it/

BUTLER, R. (1988). Enhancing and Understanding Intrinsic Motivation: The Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance. Br. J. educ. Psychol., 58, 1-14, 1988

VARK: Prithish Kumar IJ, M. SA. Understanding Your Student: Using the VARK Model. J Postgrad Med. 2014 Apr-Jun;60(2):183-6. doi: 10.4103/0022-3859.132337. PMID: 24823519.
https://www.jpgmonline.com/article.asp?issn=0022-3859;year=2014;volume=60;issue=2;spage=183;epage=186;aulast=Prithishkumar