Specifications Grading primarily bases grade on:

1. the *quantity* of work a student chooses to do. Higher grades require more work.
2. as well as on the *quality* of work. Higher grades require more sophisticated work.
INTRODUCTION
I want to give you some context for what I teach so that you can begin thinking about how all or part of it can apply to your respective fields. I teach courses a four-semester sequence in music theory and aural skills, music analysis, composition, and orchestration.
This area of music deals with both “left” and “right” brain functions: the logical side considers matters of psychoacoustics, mathematics, physics, and music cognition; the creative side considers matters of expression, creativity, artistry, and meaning. So I hope you’ll find something of value here whether you teach chemistry or dance.
The model I want to put forward is that of Specifications (or Specs) Grading. This model by no means originated with me; I'll point you to some sources shortly. But I was dissatisfied with the grading system I had been using and discovered this during the summer of 2017, and it completely changed the way I approach grading.
THE PROBLEMS WITH TRADITIONAL GRADING SYSTEMS

• Worked against my teaching goals. (Student: “I don’t care what I learn so long as I pass.”)
• Perpetuated a “game-playing” approach to education and encouraged point-grubbing and hair-splitting, rather than learning.
• Didn’t communicate what a student is (or isn’t) learning.
• Gave too many:
  ‣ false positives (students who pass but didn’t actually learn much)
  ‣ false negatives (students who consistently learn, but often after the graded work has been turned in).

Here are some of the problems I began seeing with grading systems I had used in the past. I read a similar list to this on a blog I’ll mention in a moment, and I realized that it articulated exactly the problems I’d been noticing.
My goals are actually fairly straightforward:
1. I want students to gain mastery, complete command of the skills or concepts of the course.
2. I want them to be able to apply those skills or concepts in more complex or nuanced contexts that require them to think critically and creatively.
3. I want them to engage with the course content beyond the classroom.
The progenitor of specs gradings is Linda Nilson, of Clemson University. But Robert Talbert, a mathematics professor as Grand Valley State University, is the one who brought specs grading to my attention, and whose implementation of it I have used as a model for my own courses.
The first principle of specs grading is to reduce the complexity of assessment down to a binary rubric of satisfactory and unsatisfactory. These terms can be more useful than pass/fail. As one who has always struggled with a fear of failure, I think that the current generation of students would do well to get over their aversion to failure so that they can get on with the business of eventual success. That said, these terms tend to help them get over that hurdle because they seem to bear less finality than “pass/fail.”
**RUBRICS**

- Satisfactory
- Unsatisfactory

- Criteria for “Satisfactory” is predetermined and put in writing on the quiz/assignment.
- No point grubbing.
- No splitting hairs.

**What’s the value?**

- The criteria for what constitutes “satisfactory” is clearly articulated on the assignment, or in some general rubric that applies to all assignments of a particular type. This makes it easier to grade quickly and abundantly clear to students what the threshold of acceptance is.
- There is no need to agonize over the relative merits of partial credit or how many points this or that answer should be worth.
### RUBRICS

<table>
<thead>
<tr>
<th></th>
<th>Excellent/Exemplary/Exceeds Expect.</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>The work meets or exceeds the expectations of the assignment. Communication is clear and complete. Mastery of the concepts is evident. There are no non-trivial errors. This work could be used as a classroom example.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Meets Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>Understanding of the concepts is evident through correct work and clear, audience-appropriate explanations. Some revision or expansion is needed, but no significant gaps or errors are present. No additional instruction on the concepts is needed.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Revision Needed</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Partial understanding of the concepts is evident, but there are significant gaps that remain or significant errors present. Needs further work, revision, more review, and/or improved explanations.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Not Assessable</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Not enough information present in the work to determine whether there is understanding of the concepts. The work is fragmentary or contains significant omissions/errors. Or, there are too many issues to justify correcting each one.</td>
</tr>
</tbody>
</table>

Another rubric I use, à la Robert Talbert, is the more nuanced EMRN rubric, in which E and M are satisfactory marks, and R and N are unsatisfactory. This bifurcation of the previous two-tier rubric gives students a little more information on how satisfactory or unsatisfactory their work is. These are the general descriptions I give in my syllabus, but in practice, I tailor these descriptions to the assignment type so students know what I’m expecting.
A shorthand way of looking at these would look like this. If you were to make a bell curve of student trends, about 80% would fall in the middle (landing in the green or orange categories), and the remaining 20% would land at the extremes. That said, I can anecdotally attest that the trend I’m seeing among students now in their 2nd or 3rd semester using this system, I’m handing back fewer Ns and more Es than I did the first semester that I implemented specs grading.
**Grading by Bundles**

<table>
<thead>
<tr>
<th>GRADE CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A students have not only understood but mastered the concepts, applications, and theory in the course, shown by a wide range of evidence. They are consistently actively engaged with the course by preparing for and being a leader in class activities and in office hours. They are ready to move on to further coursework without reservations.</td>
</tr>
<tr>
<td>B students have shown evidence of strong understanding of the basic skills and applications of the course, with some evidence of mastery on some topics. They are consistently ready to work in class and are actively engaged in class activities. They are ready to move on to further coursework.</td>
</tr>
<tr>
<td>C students have attained the baseline level of competency required by the course, but do not have consistent evidence of mastery or strong understanding. They may take further coursework but with some caution.</td>
</tr>
<tr>
<td>D students have not attained the baseline level of competency required by the course, although there is some evidence of understanding. They are not yet ready to move on to further coursework in this area.</td>
</tr>
<tr>
<td>F A student with an F has not shown evidence of minimal understanding of the course concepts.</td>
</tr>
</tbody>
</table>

The second principle of specs grading is the notion of assigning grades by the quantity and/or quality of that work students complete. I actually put these descriptions in my syllabus because I want to pull students’ attention away from numbers and percentages. That’s not what matters to me—what matters is their mastery of the material and their preparedness for what comes next.
### Grading by Bundles

<table>
<thead>
<tr>
<th>Achievement</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Halfling (D)</strong></td>
<td>2100 XP + 12 Learning Targets. (No Application Challenge required)</td>
</tr>
<tr>
<td><strong>Ranger (C)</strong></td>
<td>Achieve Halfling level + <strong>300</strong> more XP (total of 2400 XP) + 3 more Learning Targets + 1 Application Challenge</td>
</tr>
<tr>
<td><strong>Elf Lord (B)</strong></td>
<td>Achieve Ranger level + <strong>300</strong> more XP (total of 2700 XP) + 2 more Learning Targets + 1 more Application Challenge</td>
</tr>
<tr>
<td><strong>Istar (A)</strong></td>
<td>Achieve Elf Lord level + <strong>300</strong> more XP (total of 3000 XP) + 2 more Learning Targets + an “E” mark on an Application Challenge</td>
</tr>
</tbody>
</table>

Students earn those grades by completed bundles of assignments at a particular level of quality. I also try to make it more fun by tracking different levels of achievement with a theme. This semester’s theme is *The Lord of the Rings*. For this course I offer 20 learning targets. Students don’t have to do all of them to pass the class—in fact they don’t have to complete all of them to earn an A. They have a degree of choice over how they demonstrate their understanding. I’ll get into learning targets, XP, and application challenges later.
So the advantages (and I put this into my syllabus as well when trying to sell the idea to students) are as follows:

• Students can track their progress.
• They (and I) can easily identify their strengths and weaknesses.
• Their grades are based on concrete evidence of learning.
• Students to some extent choose how and when his knowledge is evaluated.
• Students have the opportunity to fail and learn from mistakes. (And I do consider it an opportunity).
• Students are rewarded for their mastery.
• Students know on the very first day how much work is required to earn the grade they want.
HOW I DO SPECS GRADING:
A DOUBLE-BARRELED APPROACH
I take two approaches to specs grading, based on whether the course tends to be more:

- Concept-based, or
- Skills-based

There are many other ways to organize, and there is even a spectrum between these two assumed poles. None of my classes are purely one approach or the other, but this is just a helpful way to organize my approach.
• For every course in the theory sequence, I create 20 learning targets which are positive “I can” statements. Each learning target has a quiz that assesses the student’s mastery of that concept (as well as sometimes retaining and integrating earlier concepts).

• So a quiz for Learning Target H4 might look like this:
Learning Target Quiz H4 (v2)

Mastery of basic concepts

APPROACH 1

1. The title identifies the learning target and which version it is. Students who earn “Unsatisfactory” may reattempt the same learning target at a future time with a different but similar quiz.
2. The learning target’s “I can…” statement is stated to avoid any confusion as to what the goal of this quiz is.
3. When appropriate, I include a reference to media. I give students access to such resources ahead of time and during the quiz. You could expand this to include a reading, an online video—sometimes I give them a sheet music ahead of time and tell them to analyze it as much as they can before the quiz.
4. At the bottom, I include the specifications for what constitutes “satisfactory” on this quiz. The entire quiz should take a well-prepared student about 15 minutes or less to complete.
In the case of a learning target like C6, the student won't have time to complete it if I give them all the material at the time of the quiz. Instead, I give them sheet music ahead of time with instruction to analyze it as completely as they can. Then I might give them a quiz like this...
C6: I can identify and analyze, through listening and score study, all the elements of Baroque fugue, using standard terminology.

Examine accompanying score as you listen (Spotify) to J. S. Bach’s B-flat major “Fugue No. 21,” from The Well-Tempered Clavier, Book I, BWV 866.

A. Large-Scale Divisions: Give the following divisions by measure number and beat (i.e., measure 3, beat 2 would be written as “m. 3.2.”).

Exposition
Development
Closing

B. Key Centers: Give the important key centers that begin in each of the following measures:

m. 1  m. 5  m. 22  m. 26  m. 37

C. Motivic Material: On the blank staff below, notate the first complete presentation of the subject (including appropriate clef and key signature).

D. Elements: Locate an example of each of the following elements frequently found in fugues.

This fugue's exposition is unusual in that it lacks a ________.

An entry group occurs in mm. ________.

A(n) ____________ occurs in mm. 30.2–37.1

A single entry occurs in mm. ________

Criteria for Satisfactory mark: Answers are complete, per the instructions. Notated subject is complete, including any necessary accidentals. No more than one error per section.
<table>
<thead>
<tr>
<th>HOMEWORK ASSIGNMENTS</th>
<th>OTHER ACTIVITIES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>50XP</strong> Each practice assignment checked late.</td>
<td><strong>50XP</strong> Complete Course Survey</td>
</tr>
<tr>
<td><strong>100XP</strong> Each practice assignment checked before class due time.</td>
<td><strong>100XP</strong> No absences or tardies</td>
</tr>
<tr>
<td></td>
<td><strong>50XP</strong> Find an online video explaining one of the semester's topics</td>
</tr>
<tr>
<td></td>
<td><strong>100XP</strong> Earn &quot;Satisfactory&quot; on all 20 LTQs</td>
</tr>
<tr>
<td></td>
<td><strong>50XP</strong> Explain a concept or demonstrate a skill to your peers in class</td>
</tr>
<tr>
<td></td>
<td><strong>100XP</strong> Find an example of one of this semester's topics in your own music and share it in class.</td>
</tr>
<tr>
<td></td>
<td><strong>50XP</strong> Write a very brief (1-page) reflection on what you have learned this term and how you can apply it going forward</td>
</tr>
<tr>
<td></td>
<td><strong>100XP</strong> Pass an LTQ on the first attempt</td>
</tr>
<tr>
<td></td>
<td><strong>50XP</strong> Complete Course Agreement</td>
</tr>
<tr>
<td></td>
<td><strong>200XP</strong> Pass 5 LTQs on the first attempt</td>
</tr>
<tr>
<td></td>
<td><strong>50XP</strong> Submit photos of three study groups you attended</td>
</tr>
<tr>
<td></td>
<td><strong>300XP</strong> Pass 10 LTQs on the first attempt</td>
</tr>
<tr>
<td></td>
<td><strong>50XP</strong> Submit a study/completion plan and timeline for an Application Challenge</td>
</tr>
<tr>
<td></td>
<td><strong>500XP</strong> Pass 15 LTQs on the first attempt</td>
</tr>
</tbody>
</table>

I measure engagement with course content outside of class by inviting students to rank up experience points. (Enthusiasts of role-playing games or video games will be familiar with this concept.) One of the things that has saved me tons of time is to not grade homework. So I have an answer key for each homework assignment, and students who come in and check their answers against mine get double the XP that those who check it late, whether or not they got it all right. This saves me time and helps students wrestle with their errors and generate questions.
• Application Challenges are more substantial, in-depth assignments that integrate a constellation of skills or concepts all at once, and they require a higher degree of critical thinking and creative solution-finding. I mark these with feedback for revision and the EMRN rubric.
• Students are given a choice of applications challenges, and choices within each challenge to help give them greater ownership over their endeavor. They have to find an explanation for someone else’s music, or solve musical problems themselves through composition.

<table>
<thead>
<tr>
<th>Option 1: Analysis</th>
<th>Option 2: Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Choose one of nineteen possible fugues</td>
<td>1. Choose:</td>
</tr>
<tr>
<td>2. Analyze, graph, and present findings</td>
<td>2. Compose according to instructed criteria</td>
</tr>
</tbody>
</table>

The first fugue subject is in D major, common 6me. Rhythm for first measure is 8th rest followed by 7 8th notes, second measure is one quarter note. Scale degrees: 5 3 1 6 4 2 7 8.

The second fugue subject is in G minor, common 6me. Rhythm is consistent quarter notes. Scale degrees: 5 5 3 4 5 4 3 2 1.

Both have prominent dominants, which makes for a lively discussion of real vs. tonal answers.
Here are examples of how I tailor the EMRN rubric to the type of assignment.

- In an analytical challenge, notice the differences between E-level and M-level. Both expect a clear and complete analysis, but E-level is also thoughtful. M-level contains no serious errors; E-level includes that plus takes a more nuanced approach to ambiguous passages. Diagrams or figures are adequate at M-level, versus exceptionally transparent and serve to clearly summarize at E-level.

- Similar differences can be seen in the composition challenge. Overall, what distinguishes E-level from M-level is that the final result is professional in tone or character.
So perhaps another lens through which we could understand these two approaches is not concept- vs. skill-based content, but through the number and kind of hurdles students must overcome.

- In the first approach, I cover many concepts (20 in all) that only need to be passed once. Thus students earn higher grades by jumping over more hurdles—i.e. quantity of hurdles (the quality of which are all reasonably high).
- Approach two, which I use in my skill-heavy courses, involves a smaller collection of skills that we evaluate many times at increasing levels of difficulty. Thus students earn higher grades by jumping higher hurdles, and quality becomes a more important factor when assigning marks.
In my Aural Training courses, I assign

- 10 assignments that externalize music (basically, notated music on the page and accurately turning it into musical sound), as well as...
- 10 assignments that internalize music (basically, listening to music sound and internalizing it such that they can notate it on the page). These assignments are prepared at the leisure of the student and can be turned in when they are ready (within certain limits). Notice that the skills are actually very similar: P1 and L1 are both about accuracy of pitch; P2 and L2 are both concerned with accuracy of rhythm, etc. These assignments are marked with the EMRN rubric and can be reattempted as much as needed. I supply students with feedback, but never answers.
- The acuity targets are skills that need to be perfected in real time, and these I quiz over the course of the term.
A prepared performance assignment might look like this, where they are given a list of exercises to practice and perfect during the week. Since I teach two such courses per semester, I don’t have time to meet with each student individually, so I outsource this to peer coaches, older students who have done well in the class and who listen to individuals’ progress the assignment. The coach requests a random selection from the list, then tracks how many errors and which kinds for each exercise, and hands off the record to me. I then assign EMRN, based on the number and type of errors and the relative difficulty of the chosen exercises.
Each Performance Assignment contains the exercises to be prepared and any special instructions or grading criteria that pertain to it. Your preparation on each assignment is submitted by meeting one-on-one with your musicianship coach. When you meet with your coach, he/she will give feedback on how to improve. I will grade the assignment based on conversation with and feedback from the coach (as well as in light of how you are doing in class).

**EVALUATION**

All Performance Assignments are evaluated using the EMPN rubric outlined in your syllabus. Below, you’ll find more specific descriptions of what those marks mean in terms of these assignments.

- **E:** Performance is not only accurate, but confident, well-prepared, and musically attentive. Pitches and rhythms are precise. Conducting pattern is aligned with performance, and hand direction reflects the beat. Phrasing, dynamic, articulation, and tempo marks, observed. Solfège is accurate. No significant hesitations or restarts.

- **M:** Performance is accurate and there are no systematic errors. Pitches and rhythms are mostly accurate. Conducting pattern is generally aligned with performance, and hand direction reflects the beat. Phrasing, dynamic, articulation, and tempo marks, observed. Solfège is accurate. Some slight hesitation but no restarts.

- **P:** Performance contains some significant errors that suggest more practice and preparation is needed. Pitches and rhythms are sometimes accurate. Conducting pattern is inconsistent and often gets in the way of performance. Phrasing, dynamic, articulation, and tempo marks, are largely ignored. Solfège is inconsistent. There is at least one significant hesitation or restart.

- **N:** Performance contains several significant errors; or there are portions of the assignment that could not be performed. Pitches and rhythms are mostly inaccurate. Conducting pattern is haphazard/non-existent. Phrasing, dynamic, articulation, and tempo marks seem non-existent. Solfège is haphazard/non-existent. There are multiple significant hesitations or restarts.

Copyright ©2017 by Enoch S. A. Jacobus and licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

I use this rubric as my guide when doing so, and this sheet is also given to the students and the coaches, so that everyone knows what we’re aiming for.
Transcription assignments like this one are evaluated on how accurately the student has heard and transcribed pitch, rhythm, notational accuracy/clarity, and accuracy of following instructions/answering follow-up questions on the excerpts. I use a detailed EMRN breakdown for these assignment as well, much like those I just showed.
Aural Quiz III.1

I. Seventh-Chord Identification
In the following blanks identify the seventh chords you hear as MM (major-major), Mm (major-minor), mm (minor-minor), ø (half-diminished), or o (fully-diminished). On the second hearing, you may softly
arpeggiate the chord to yourself. (Each example will be played 2x.)

1. ________    2. ________    3. ________    4. ________    5. ________

Criteria for Satisfactory mark: MS must clearly delineate upper and lower case. No more than one error.

II. Harmonic Progression
Using Roman numerals and inversion labels, mark the following harmonic progression. You may use the
blank line below as a guide for the Do-Ti test. (The progression will be played 3x.)

Chords: _____                        ______                        ______                    _______                    _______

Criteria for Satisfactory mark: Roman numerals must be clear, reflect the correct chord quality, and include
any necessary inversion symbols. No more than one error in Roman numeral identification. No more than 2
errors in inversion identification.

III. Binary Form Identification
Indicate whether the form heard is simple or rounded binary, and whether it is sectional or continuous.
(The excerpt will be played 2x)

Circle one:  Sectional / Continuous

Circle one:  Simple / Rounded

Criteria for Satisfactory mark: No errors present

IV. Error Detection (Ottman 8th ed. 14.48)
The following duet from your singing assignments will be played with errors in pitch, rhythm,
ariculation, and dynamics. For each error, on your score circle the note(s) affected by the error, label the
type of error that occurred (P, R, A, or D), and indicate what was actually played. (The incorrect version
will be played 3x.)

Criteria for Satisfactory mark: Detected errors must include circle/bracket of affected notes, error type
label, and indication of what was incorrect. No more than one missed or misidentified/mislabeled error.

On this aural acuity quiz you can see four of the five acuity learning targets being evaluated. Each section has its own criteria for satisfactory completion.
Here's something of a point-by-point comparison between traditional grading and specs grading.
## Specs Grading Advantages

- Can readily accommodate quantitative assessments
- Far less grading (and laboring over the minutia of point values)
- Fewer conflicts with students*
- Spend less time evaluating qualitative tasks
- Greater motivation for meaningful learning through quality work
- Students more likely to learn from errors and corrective feedback
- Less stress and likelihood for cheating*
- Improved student performance over time (through stronger models)
- Reasonably high standards
- Early warning of student weaknesses

## Specs Grading Disadvantages

- Unfamiliarity
- Requires comprehensive plan for entire course
- Need for clearly defined, evaluative criteria and models of acceptable performances
- Likely to require more feedback


And here’s a list of pros and cons for specs grading.
CONSIDERATIONS FOR CONVERSION

1. Search the web for examples in your discipline. There are many out there in a variety of subjects and fields.

2. Start with a partial conversion (one or two assignments or bundles)


4. Bounce ideas around. I hope you will view me as a resource.