

Skills-Based Grading: A novel approach to teaching formal semantics

Maura O’Leary & Richard Stockwell*

Abstract. This paper reports an implementation of ‘Skills-Based Grading’ (SBG) in a formal semantics course. In traditional grading, every part of every assignment contributes to the final grade. Students are required to progress along a uniform timeline, with partial credit as a safety net. In SBG, by contrast, the course is composed of skills. Students are given multiple opportunities to demonstrate mastery in each skill, but full proficiency is required to gain credit. Zuraw et al. (2019) pioneered the use of SBG in linguistics for phonetics and phonology. SBG is known to work well for skills that require algorithmic approaches to arrive at inarguably correct answers. In applying SBG to semantics, we show that it is just as effective for more abstract and philosophical skills. Based on survey and grade data, we substantiate claims that SBG improves student learning, encourages more effective study, lowers student stress, and achieves more equitable outcomes. Since this paper reports our first use of SBG, we conclude with some reflections on improvements for the future.

Keywords. Skills-Based Grading; semantics; pedagogy

1. Outline. This paper reports an implementation of ‘Skills-Based Grading’ (SBG) in a formal semantics course. Section 2 defines SBG in opposition to traditional grading. Section 3 describes our application of SBG to semantics, which is then evaluated in section 4. Section 5 offers some concluding remarks and reflections.

2. Traditional vs. Skills-Based Grading. A typical course in the North American education system is composed of a series of assessments – homework assignments, quizzes, midterm and final exams – each of which contributes a proportion of the final grade. This grading structure requires students to progress along a uniform timeline: every part of every assessment counts;¹ lost points cannot be recovered. Perhaps to compensate for this inflexibility, partial credit is commonly awarded for answers that show partial understanding of the material. Call a system of this kind ‘Traditional Grading’ (TG).

In ‘Skills-Based Grading’ (SBG), by contrast, a course is composed of skills. The list of skills that make up the course is presented to students at the outset. Each exercise on an assessment then provides an opportunity to demonstrate proficiency in one or more skill. An opportunity is offered in the week a skill is taught and at least once subsequently. Attempts are

* Thanks first and foremost to our students for their patience and feedback. Thanks also to Douglas Ezra Morrison for help handling the data; Kie Zuraw and Franny Brogan for their advice and encouragement with Skills-Based Grading; Jessica Rett for her advice throughout; Dylan Bumford, Ethan Poole, Yael Sharvit, Giuseppina Silvestri and Jos Tellings, for whose semantics courses we have been Teaching Assistants; three anonymous LSA reviewers for their feedback, and for rating our abstract to place third in the Student Abstract Award; and audiences at UCLA and the LSA. Authors: Maura O’Leary, University of California, Los Angeles (mauraoleary@ucla.edu) & Richard Stockwell, Christ Church, University of Oxford (richard.stockwell@chch.ox.ac.uk).

¹ There are ways to mitigate this aspect of traditional grading; for example, by discounting a student’s lowest-scoring homework. However, such mitigations are liable to both under- and over-perform. On the one hand, they are insufficient for late blooming students, or those who suffer setbacks across multiple weeks. On the other hand, they are overly generous to students who perform consistently throughout the course, and are then free to (ab)use the mitigation to make no attempt at the last homework. For related discussion, see sections 4.4 and 5.

graded as “not yet proficient”, “approaching proficiency”, or “proficient”. Only “proficient” answers count towards final grades; “not yet” and “approaching” are for guidance only. While there is no penalty for incorrect answers, only fully correct answers are marked “proficient” – there is no partial credit.

To compare directly, where a TG course is composed of each point on each assignment, an SBG course is composed of skills. On the one hand, SBG is more generous than TG: students are granted multiple opportunities to demonstrate proficiency in each skill. On the other hand, SBG is more rigorous: students are required to demonstrate full proficiency in a skill, without the safety net of partial credit. We found SBG’s balance between generosity and rigor to be one that more effectively drives and measures student learning. These claims are substantiated by the evaluation in section 4. The next section describes our application of SBG to semantics.

3. Applying Skills-Based Grading to Semantics. Implementations of grading schemes along the lines of SBG have been reported in a variety of secondary and higher education settings; for example, standards based grading (Schimmer 2016), mastery based grading (Armstrong & Pet-Armstrong 2003, Brackett & Reuning 1999), and specifications grading (Nilson 2015).

Zuraw et al. (2019) pioneered the use of SBG in linguistics for phonetics and phonology. We report an application of SBG to formal semantics. Further to mathematical and algorithmic skills, where questions have a ‘right’ answer, semantics encompasses more abstract and philosophical skills, with topics including possible worlds, vagueness, and pragmatics. Semantics also involves skills requiring knowledge as a language user; creating original examples, for instance, or identifying meaning patterns such as contradiction and ambiguity. In implementing SBG for semantics, we demonstrate that SBG is applicable to these more abstract skills.

We implemented SBG in an introductory formal semantics course at the University of California, Los Angeles (UCLA Semantics 1, Ling 120C). The 29 students had taken two prerequisite courses: an introductory linguistics course for intending majors and introductory syntax. The course was taught online in a six-week summer session, with familiar limitations. In particular, the very intensive timeframe meant learning was very compressed, with no time for, e.g., a midterm exam.

The course was composed of 47 skills, organized into 12 ‘skill groups’. Ten of the skill groups corresponded to topics in the course material; the other two were for active participation and analytical thinking, discussed later. The ten course topics were: ambiguity, word meaning, sense and reference, sentence relationships, propositional logic, set theory, negative polarity items and tonicity, predicate logic, model theory, and lambda calculus. Students started out from a grade of 20%. This baseline reflected our judgement of the level of skill proficiency required to align with traditional grade boundaries. For the most part, one demonstration of a skill contributed one percentage point of a student’s final grade. Some skills required multiple demonstrations, as we will see with reference to the following examples.

The skills comprising the fourth skill group, ‘Sentence Relationships’, are listed in Table 1:

Skill identifier	Sentence Relationship skills	Demonstrations required
4.1	Identify implicature and maxim	4
4.2	Test for asymmetric entailment	1
4.3	Test for mutual entailment	1
4.4	Test for presupposition	1
4.5	Test for contradiction	1

Table 1. The Sentence Relationships skill group

Notice that skill 4.1 requires four demonstrations – one per Gricean maxim. This skill was not further decomposed, so as to avoid giving away the answer to a particular question. An example question providing an opportunity for one demonstration of proficiency in skill 4.1 is given in Figure 1:

1. Identify an implicature that A might interpret from B's utterance in the conversation below. Then state which maxim was flouted to communicate this implicature.

A: *Would you recommend Hama for this tenure-track professor position?*

B: *She has excellent handwriting.*

Skill 4.1: Identify implicature & maxim

☐ not yet proficient

☐ approaching proficiency

☐ proficient

Figure 1. A question testing for a Sentence Relationships skill

The question is labelled with its skill, along with a box for the grader to indicate the demonstrated level of proficiency. A “proficient” answer would earn one percentage point of the final grade; “not yet” and “approaching” have no effect on grades, being for guidance only.

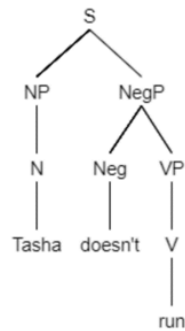
For a second example, the ‘Lambda Calculus’ skill group is listed in Table 2:

Skill identifier	Lambda Calculus skills	Demonstrations required
10.1	Calculate semantic type	3
10.2	Give denotation (simple: proper name or predicate)	1
10.3	Give denotation (complex)	3
10.4	Do derivation (simple: only proper names and predicates)	1
10.5	Do derivation (complex)	2

Table 2. The Lambda Calculus skill group

Again, some of the skills require multiple demonstrations, with each demonstration contributing one percentage point to the final grade. An example question for skill 10.1 is given in Figure 2:

2. In class, we saw how sentence-external negation might be formalized. Your task here is to figure out the semantic type of sentence internal negation (“doesn’t”) using the tree below.



Skill 10.1: Calculate semantic type

- ☐ not yet proficient
- ☐ approaching proficiency
- ☐ proficient

The semantic type of sentential “doesn’t” is _____

Figure 2. A question testing for a Lambda Calculus skill

Demonstrating full proficiency across the ten skill groups corresponding to topics in the course material scored 86% – a solid B grade. In order to achieve an A grade, students had to show some proficiency in the remaining two skill groups. Four points were available for active class participation under ‘Explain’. This skill group required participation on an online forum and in class, including presenting a real-world example illustrating a semantic concept. We found that SBG was straightforwardly applicable to participation, allowing the whole grade to be determined in SBG style.

The last skill group was ‘Advanced Analysis’. Up to ten points were available under this heading for proficient answers to questions that required students to think beyond the lecture material, evaluating theories and applying them to further data. An example advanced analysis question is given in Figure 3:

3. *Context: Your airline ticket includes a single drink.*
‘The flight attendant will give you tea or he will give you coffee.’

- (a) What does the \vee operator predict for this sentence when both ‘he will give you tea’ and ‘he will give you coffee’ are true?
- (b) Given the context, do you think that it is possible that both disjuncts are true?
- (c) Answer one of the following prompts.
 - How can you reconcile this difference, based on what we have learned in class?
 - Provide a formula in propositional logical, using the operators we have learned, that matches your intuitions about what is true in this context. (That is, the entire truth table for that formula should match your intuitions.)

Advanced Analysis

- ☐ not yet proficient
- ☐ approaching proficiency
- ☐ proficient

Figure 3. A question testing for Advanced Analysis

With these two skill groups beyond course topics, there was thus a meaningful distinction between a B grade and an A.²

Having described our implementation of SBG for semantics, the next section evaluates its effectiveness.

4. Evaluating Skills-Based Grading for Semantics. In order to evaluate the effectiveness of using SBG in teaching semantics, we collected quantitative and qualitative data as the course progressed. Further to numerical data from weekly grade updates, we conducted surveys at the midpoint and end of the course.³

The qualitative feedback was broadly positive; for example:

- (1) a. “This is the most transparent grading scheme that I’ve experienced!”
- b. “It was really great! I wish all my classes did this.”
- c. “I think a student who succeeds academically in courses with traditional grading methods won’t have any trouble succeeding in a course with skills based grading, and someone who struggles with traditional grading may find skills based grading easier.”

Further, students professed to liking and understanding SBG, as shown in Figures 4 and 5:

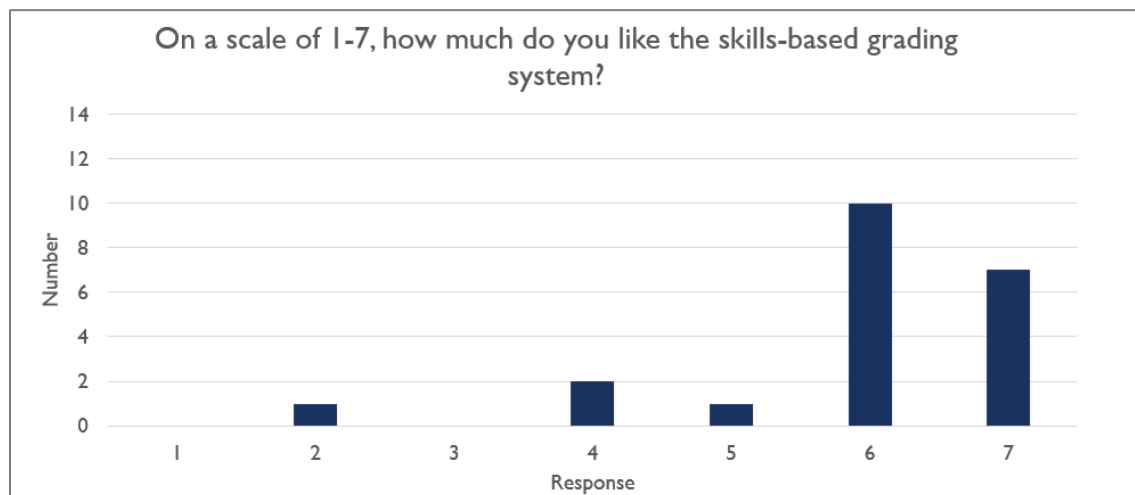


Figure 4. Student liking of Skills-Based Grading

² Zuraw et al. (2019) likewise had a meaningful distinction between B and A grades, though via more complex formulae – see their (1)-(4), pp. 412-5. In our system, crossing the threshold was simply additive, which we hoped would be easier to understand.

³ Our methods were approved under UCLA IRB #20-001297. Students were rewarded with one percentage point of extra credit for completing each survey.

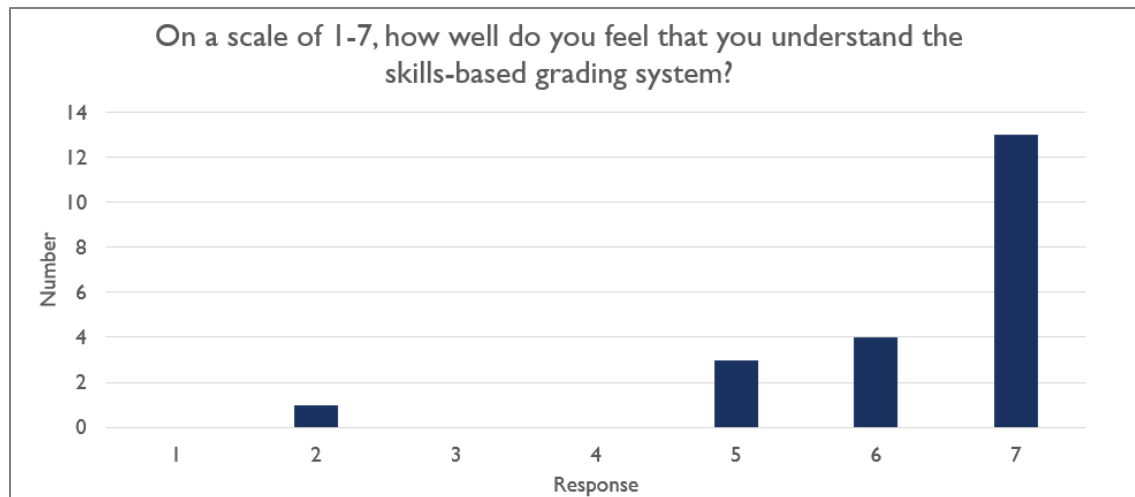


Figure 5. Student understanding of Skills-Based Grading

The rest of this section uses our grade and survey data to evaluate our application of SBG to semantics. We proceed with reference to four benefits that have been claimed for SBG in the literature: (i) improving student learning, (ii) encouraging more effective study, (iii) lowering student stress, and (iv) achieving more equitable outcomes. We also (v) assess the applicability of SBG to the more abstract parts of semantics, before section 5 offers some concluding reflections.

4.1. IMPROVED STUDENT LEARNING. SBG’s lack of partial credit offers a better measurement of student learning. SBG assesses complete mastery rather than partial understanding, as answers need to be completely correct in order to earn credit. While partial credit is often used in TG to encourage students to attempt an exercise, it can also incentivize guessing (Nilson 2015). SBG instead encourages attempts by removing the potential for a grade penalty entirely. In an SBG system, students are allowed to re-attempt skills as many times as they are offered until they reach a level that demonstrates complete proficiency, with no penalty for incorrect attempts. While students may still guess at an answer during these attempts, the chance of them guessing an answer that is completely correct is low, and thus the guesswork will likely not contribute to their grade. We found in using TG in previous courses that there is a strong tendency to give partial credit for merely attempting a question, which leads to the potential for a passing grade from a student who largely guessed their way through the course.

The all-or-nothing nature of SBG worked just as well for the more abstract topics in semantics. We did not have to award partial credit for answers that were overly vague or that used key terminology incorrectly. For instance, if a student confused strict and lax with respect to dictionary definitions, TG might still award partial credit for any (mis)labeled evidence provided. In SBG, no credit is awarded until the student achieves full understanding.

The lack of partial credit also encourages students to move away from a ‘remember and regurgitate’ style of study and towards a lasting understanding of the material (Buckmiller, et al. 2017). If a skill is not mastered in the first week it was taught, students must continue to work on developing that skill throughout the course, until they can demonstrate that they have learned it fully. Students saw this value in the system:

- (2) “I think it is an effective and accurate way to assess knowledge”

One might worry that if students can learn a skill and immediately demonstrate it, then the fact that they do not need to return to it will allow them to forget it. Generally, in a TG course, this problem is combated with cumulative – and mandatory – midterms and finals. (It should be noted that mandatory cumulative tests are not incompatible with this system, as will be discussed in section 5.) However, even without a mandatory final exam, SBG courses can be designed in such a way that the material is not immediately forgotten. One simple method is to require more than one demonstration of a given skill, and to not immediately offer enough exercises to meet the maximum number of demonstrations. Delaying additional opportunities for a specific skill can force students to revisit material at a time defined by the instructor, which is useful when a complex skill taught late in the course builds on an early skill.

More organically, instructors can rely solely on the ‘building’ nature of skills throughout the course. We found that this was enough to keep students engaged with early skills, even without explicitly requiring students to attempt early skills over again. Even when an easy skill was mastered immediately, it was still in use throughout the course – for instance, the tools introduced in propositional logic (an early set of skills) were used in predicate logic (a mid-course topic), before predicate logic was in turn used in lambda calculus near the end of the course. Likewise, set theory (another early-course topic) was used in the discussion of predicate logic, lambda calculus, NPIs, monotonicity, and so on. Because early material was used in later weeks, we found that students rarely learned a skill and then left it behind. It also seems likely that applying earlier skills to later skills may help some students to grasp ‘building block’ skills that they may not have understood initially. In TG, students must demonstrate skills in the week they are taught; if seeing propositional logic used in the context of predicate logic finally causes propositional logic to ‘click’ for a student, it will be too late to recover the points for propositional logic. SBG, by contrast, allows students to earn credit for early skills in the same week that they earn credit for advanced skills.

SBG helps students to be more aware of the skills that make up the course material. This in turn enables them to see how skills build on one another, as well as to track their own growth as they gain new skills (Zuraw et al. 2019). One student wrote:

- (3) “I think there’s a lot of value in this grading system and I appreciate how it makes you think more about each individual constituent skill that comprises understanding the course (as well as understanding semantics), and I think my understanding of the material ... is better than what my understanding of the material of a different given course would be”

Students also appreciated the self-paced nature of SBG. When asked about the aspects of SBG that worked well for them, a representative response was:

- (4) “Self-paced learning and not losing points for not mastering a skill at one point in time”

While the lack of a forced timeline was preferred by many of our students, some were concerned about being able to delay earning points:

- (5) a. “Knowing that you will have other opportunities to demonstrate a skill has made it incredibly difficult to stay on track and motivated with the class. Instead of taking the time to get it right now, I put it off, I don’t go to office hours, and I leave it up to the future to figure out the skills I haven’t gotten yet. This grading system promotes laziness, for me at least.”
- b. “I feared getting complacent and telling myself “oh I can master the skills/get the points later”, which I feel would have caused an insurmountable amount of work towards the end of the course.”

On the whole, our class did not behave in the way suggested in (5). Rather, grades increased very steadily throughout the course, as shown in Figure 6:

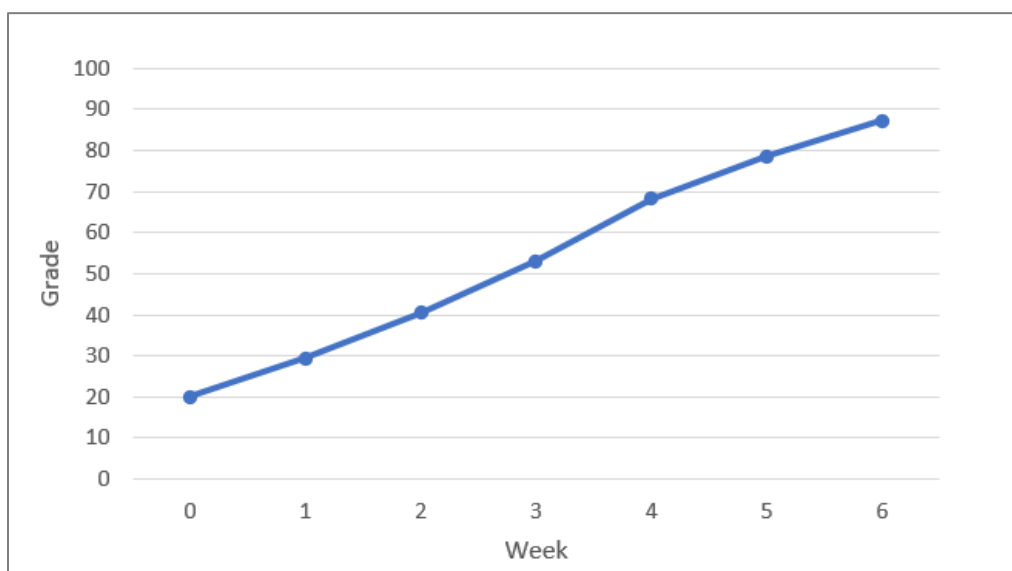


Figure 6. Average grades by week

In future iterations of this course, we intend to include this graph in the syllabus, to encourage students to aim for the same steady growth. After showing students this data from the beginning of the course, we believe they would have much less uncertainty about how quickly they might want to earn their points. As will be discussed in section 4.3, we also intend to give students, from the beginning, a more accurate idea of which assignments will offer attempts at which skills, giving students the information they need to plan their learning throughout the course.

Providing the graph in Figure 6 might also provide comfort to students who were unsettled by the unfamiliar additive nature of SBG grades:

- (6) a. “your total grade for that whole time is shown as a part of the total points possible for the whole class. Therefore, near the beginning of the class it will report you will have a terrible grade”
- b. “Although my grade so far is listed, I feel as though there is still some confusion about where it will be eventually.”
- c. “trying to guess what my end grade would be was difficult”

In general, students at North American universities are used to the subtractive nature of TG, wherein grades start at 100% and points are lost for any incorrect or absent answers. SBG, by

contrast, starts the students at a low grade (20% in our course) and then adds points as they are earned. This means that at the beginning of the term, every student's grade showed 20% out of 100; grades then averaged around 30% at the end of week 1, 40% at the end of week 2, and so on. To students who are used to a subtractive traditional grading system, 40% at any point during the term may be conflated with 40% in a TG course – an incurable, failing grade. To remedy these preconceived notions about what 40%, etc., means for your final grade, instructors will need to regularly reassure the students about the SBG system and where their current grade might land them if they continue to earn points at the same pace. Even though a prominent feature of SBG is that it does not require students to increase their grade steadily throughout the course (e.g. a student could in principle have earned 100% from our final exam alone), we found that most of our students exhibited steady growth nonetheless.

Another option to assuage student concerns would be to explicitly communicate how many points have been available thus far. For instance, by the end of week one, students in our course could have had a maximum grade of 34%. While this information might be appreciated by the students, we do not believe that it would be beneficial for students to see their grade as a percentage of points available so far; this would regress the system back towards TG, where grades can decrease week to week.

4.2. EFFECTIVE STUDY. A second benefit of SBG is that it allows a more effective approach to studying (Buckmiller et al. 2017, Zuraw et al. 2019). Once students are proficient in a skill, there is no need to demonstrate that skill again, which means that students can avoid repeated study and assessment of skills they have mastered. This allows them to focus their efforts on skills they still need to work on. Representative comments were:

- (7) a. “I can focus on things I need more practice on rather than things I already mastered.”
- b. “I like how the system shows what you do/don’t understand so you can allocate more time into working on the parts you have troubles with.”

SBG helps to make students aware of which skills they have down and which skills they need to continue working on, and then allows them the time to more effectively act on that knowledge. SBG also incentivizes reattempting skills that were initially failed; because a failure does not affect the grade in any lasting way, students can treat unsuccessful attempts as a learning opportunity rather than a failure. Students are therefore more likely to adopt a growth mindset (Dweck 2008) and treat mastery of a skill as something to be worked towards (Zuraw et al. 2019) rather than something to be tested on. Student comments reflected this:

- (8) a. “It encourages growth and a better mindset about ‘failing’, and turns ‘no’s to ‘not yet’s.”
- b. “No cost for trying”

In SBG, final grades can only be improved by any given attempt or assignment, which means that students can take the time they need to grasp more complex topics. This again stands in stark contrast to traditional, subtractive grading, where students are permanently penalized for mistakes. This was listed by our students as a benefit of SBG:

- (9) “Being able to skip questions you don’t understand at the moment and then return to them on another assignment once you understand them more as opposed to being penalized for being wrong. Once you lose points for a certain thing in traditional grading you can never recover them even if you learn the material later.”

Since many skills in an introductory semantics course build on one another, the delayed reward of SBG was especially valuable in our course. As students saw the application of building-block skills to more complex skills, they grew more likely to grasp earlier skills later in the course. SBG allowed students to revisit these fundamental skills after seeing them applied to more involved topics, and to earn credit as if they had learned the skills immediately after they were taught.

4.3. LOWER STRESS. Many instructors who have worked with systems similar to SBG have reported lowered student stress (e.g. Buckmiller et al. 2017, Zuraw et al. 2019). No specific exercise is ever required, nor will it make or break a student's final grade. Until the final there is always another chance to demonstrate a skill. This means that each individual assessment brings less stress, fear, and time pressure, which our students appreciated:

- (10) a. "...there was some relief of stress that came with the mention of multiple (and without penalty) attempts of questions on homework assignments."
- b. "The fact that I would not be penalized for not demonstrating a skill proficiently alleviates stress and encourages me to try again, which is very helpful to build resilience and motivation."
- c. "It allows me to track my progress in the class, to see what skills I need to strengthen, and it causes less stress (from fear of missing points or getting a lower grade)"

Most students felt that they did not need to stress over any particular assignment, since every opportunity would be repeated in later assignments and on the final. However, some students were still concerned about the number of opportunities they would get:

- (11) a. "If I haven't mastered a skill, it's a little bit stressful wondering how many opportunities I'll have left to demonstrate it."
- b. "it would be better if we can know how many chances that we can have for fulfilling a skill"

Students did have some indication of the number of opportunities they would get to demonstrate each skill: in the syllabus we stated that "*There will be more than one opportunity to demonstrate a skill that needs to be demonstrated once, more than two opportunities to demonstrate a skill that needs to be demonstrated twice, etc.*" By the middle of the course, we had decided that each skill would be offered a minimum of three times the number of demonstrations required to receive full credit (i.e. for a skill that needed to be demonstrated once overall, there would be at least three opportunities between the homework assignments and the final; for a skill that needed to be demonstrated twice, there would be at least six opportunities, etc).

Although we tried to remind students regularly about the number of opportunities they would get, it seems that we did not fully alleviate their concerns, as evidenced by the comments in (11). In the future, in addition to regularly reminding the students of how many attempts they will get in total, we intend to provide a spreadsheet specifying, for each assessment, which skills will minimally be offered. It is worth noting that we do not believe that the distribution of skills throughout assignments should be fully determined in advance, as we found it helpful to be able to target assignments to skills that more students were struggling with (as did Zuraw et al. 2019), which is likely to change somewhat between course offerings.

4.4. EQUITY. SBG additionally offers a more equitable system of evaluation for a diverse student body. There will always be factors that prevent students from giving their all to every assignment: illness, grief, global pandemics, and so on. In SBG, there is no penalty for missing assignments due to external setbacks, which students found comforting:

- (12) a. “I really like that my grade couldn’t go down, especially with everything going on and the class being remote it was very comforting to know nothing could harm me, whether it was missing an assignment or a question.”
- b. “Receiving second chances at receiving points when outside circumstances may have hindered my current week’s grades”

In TG, one missed homework can severely impact a student’s final grade. Some instructors try to combat this by dropping the lowest quiz or homework score, but this does not allow for students who are faced with outside factors that affect their performance in more than one homework or more than one quiz. In SBG, because no singular homework assignment *needs* to be completed, grades are not negatively impacted by a student’s inability to complete a homework assignment during difficult times. Nor do students need to complete the entirety of an assignment. Many of our students availed themselves of this opportunity and only answered a few questions on a long assignment, earning themselves a few percent of their final grade without there being a penalty for not completing the rest of the assignment.⁴

The leniency of SBG is particularly beneficial to students facing structural and institutional disadvantages. Those students may not enter the class with the same background and are more likely to experience external stressors which lead to missed classes or assignments. This was important to us, as 32% of UCLA undergraduate students are first generation students⁵, 36% have transferred from another higher education institution, 26% learned English as a second language⁶, and 35% are Pell grant recipients⁷. While listing positive aspects of SBG, one of our students noted how the system was beneficial to their classmates in this respect, especially in the context of an online class:

- (13) “Very clear expectations on what students get out of the class; easy to see strengths and weaknesses; less exam-induced stress for students in poorer living conditions and students in different timezones.”

An added bonus of this system is that there is no need for instructors to make judgment calls about what justifies extensions. Most instructors have encountered emails from students asking for deadline extensions for varying reasons and have been forced to decide whether to allow a late submission, thereby treating the student differently from the rest of the class. In SBG, there

⁴ SBG cannot fully eliminate the pressure to complete assignments along a particular timeline; for instance, skills learned in the last week of the course will need to be demonstrated on the last homework or the final exam. In a longer course than ours, this might be solved by saving the final few weeks of the term to focus on a final project or paper. During those weeks, students could catch up on skills from earlier in the course without any new skills being taught.

⁵ <https://firsttogo.ucla.edu/About/Campus-Demographics> (2016-17)

⁶ <https://www.apb.ucla.edu/Portals/90/Documents/Campus%20Stats/UGProfile18-19.pdf> (2018-19)

⁷ <https://www.ucla.edu/about/facts-and-figures> (2020-21). Pell Grants are federal grants available to low-income undergraduates in the US. The number of Pell Grant recipients at a university is often considered to be the most accurate measure of economic diversity (<https://www.usnews.com/best-colleges/rankings/national-universities/economic-diversity>).

is no need for extensions, as the students' grades will not suffer from turning in a partially completed assignment or from not turning in a particular assignment at all.

4.5. SBG IN SEMANTICS. Zuraw et al. (2019) showed that SBG is applicable to the evaluation of the concrete skills learned in phonology and phonetics courses. Our course demonstrated that SBG can be successfully adapted to semantics, in which there are concrete skills with a single right answer, but also skills that are more abstract or philosophical with infinitely many possible answers and skills that draw on knowledge of meaning as a language user; for example, coming up with your own examples or noting when two sentences contradict one another. We found SBG to work equally well with each of these skill types and did not find any significant difference across types in how quickly students moved from being taught a skill to demonstrating proficiency in it. Across concrete, abstract, and language user skills, students earned points along more or less the same timeline and mastered around the same number of skills on average by the end of the course. Concrete skills accounted for 41% of the final grade, and students earned, on average, 85% of the potential points for these skills. Abstract skills accounted for 30% of the final grade; students earned, on average, 76% of the potential points. Skills involving knowledge of meaning accounted for the remaining 29% of the final grade; students earned 84% of the potential points on these skills.

5. Concluding remarks. This paper has reported an application of SBG to semantics. Our survey data demonstrated that SBG has many benefits: improved student learning; more effective learning, encouraging students to move on from skills they have mastered; lowering student stress with additive grading, there being no way for students to permanently damage their grade; and achieving more equitable outcomes. In addition, SBG works well for a wide variety of content, including the more philosophical and language user-based aspects of semantics.

For future iterations of the course, the student feedback raised four main points we would want to address. First, we would label skills on lecture slides and handouts as they are taught, fulfilling this representative request:

(14) “matching a skill with the lecture notes/slides”

Second, we would conclude the course with a cumulative and mandatory component such as a final paper or final exam. As one student opined:

(15) “I think the final should be mandatory. It would be a good indicator of whether or not I have really understood the course content.”

A culminating component would guard against an unfortunate aspect of SBG – the opportunity for ‘early check out’. In our course, it was possible to have achieved a B grade by the end of week 5 of 6. Having done so, one student ceased participating in the course entirely. While they achieved their target grade, this student would be ill-prepared for follow-up courses, having left before seeing a λ .

Third, some students requested (16):

(16) “More homework opportunities to demonstrate skills!”

Offering more opportunities would further lower student stress. However, we believe that some time pressure is useful to encourage students to value each opportunity at a skill, and to disincentivize guessing. Striking the right balance between opportunities and pressure is a delicate matter, but sits at the heart of SBG. We would hope to fine-tune this balance over future iterations of the course.

Fourth, in composing exercises, we struggled to create multiple instances of adequately different yet consistently difficult questions testing for simple skills. Contrary to the spirit of SBG, this meant that exercises testing a particular skill often grew more difficult over time. Some students noticed this:

- (17) “I noticed that each subsequent time a skill was offered on the homework, the problem seemed harder to me ...it struck me as not very equitable towards those who struggled to understand or master a skill when the application of the skill was easier before.”

Of these four issues, the last may prove to be the most difficult to remedy. Still, we feel that the benefits of SBG overall outweigh these issues, and we both intend to use SBG again in the future.

References

- Armacost, Robert L. & Julia Pet-Armacost. 2003. Using mastery based grading to facilitate learning. *33rd Annual Frontiers in Education (FIE 2003)*. T3A-20.
<https://doi.org/10.1109/FIE.2003.1263320>.
- Brackett, Carolyn C. & Richard H. Reuning. 1999. Teaching pharmacokinetics using a student centered, modified mastery based approach. *American Journal of Pharmaceutical Education* 63(3). 272–276.
- Buckmiller, Tom, Randall Peters & Jerrod Kruse. 2017. Questioning points and percentages: Standards-based grading (SBG) in higher education. *College Teaching* 65(4), 151–157.
<https://doi.org/10.1080/87567555.2017.1302919>.
- Dweck, Carol S. 2008. *Mindset: The new psychology of success*. Random House Digital.
- Nilson, Linda B. 2015. *Specifications grading: Restoring rigor, motivating students, and saving faculty time*. Stylus Publishing.
- Schimmer, Tom. 2016. *Grading from the inside out*. Bloomington, IN: Solution Tree Press.
- Zuraw, Kie, Ann M. Aly, Isabelle Lin & Adam J. Royer. 2019. Gotta catch’em all: Skills grading in undergraduate linguistics. *Language* 95(4). e406-e427.
<https://doi.org/10.1353/lan.2019.0081>.