

## A Benchmarking System for Assessment: An Experiment in Creating More Transparency in Grading

*Boyd Fuller & Sohni Kaur\**

*Editors' Note: After analyzing a tendency for existing systems of assessment to imply something that is obviously not true – that all topics taught in a negotiation course are equally important – and realizing that typical systems of grading tend to conflate data in such a way as to obscure which particular subtopics a given student may need more help with, the authors decided to try something new. Their “benchmark” approach provides for two different kinds of benchmarks, absolute and accumulative. Absolute benchmarks determine whether a student has demonstrated understanding of certain concepts or skills, and set the bar for achievement up to a certain grade. Accumulative benchmarks are used for the higher grades, and allow a degree of student choice as to what a given student seeks to focus on. Fuller and Kaur report on initial use of this system, and find that while it requires further development to be effective in encouraging students to prioritize and excel, it already works well for distinguishing which particular concepts students have learned well, and which ones demand either more coaching, or more subtle teaching.*

### Introduction

This paper describes the use of a “benchmark” system to grade students for a basic negotiation course (Negotiation and Conflict Management) taught at the Lee Kuan Yew School of Public Policy, National

---

\* **Boyd Fuller** is the instructor of the course and an assistant professor in the Lee Kuan Yew School of Public Policy, National University of Singapore. His email address is [boyd.fuller@nus.edu.sg](mailto:boyd.fuller@nus.edu.sg). **Sohni Kaur** is a graduate student at the Lee Kuan Yew School of Public Policy, National University of Singapore, and was a Teaching Assistant for this course. Her email address is [sohni.kaur@yahoo.com](mailto:sohni.kaur@yahoo.com).

University of Singapore. This system was an experiment in creating a more transparent system for assessment, which we hoped would

- 1) clearly link students' understanding of certain skills and concepts, and sometimes their ability to apply them, to each step they would have to take to improve their grade;
- 2) help the students and the instructors track, and respond to, gaps in their understanding;
- 3) prioritize the concepts and skills in the course, and require the mastery of basic elements, before proceeding with the other elements; and
- 4) allow the students to allocate their efforts more efficiently within the course and among all of their courses for the semester.

The benchmark assessment system was a response to some shortcomings in more typical approaches to assessing students. These systems tend to aggregate students' outcomes over a number of assignments, and so lose important data about the specific skills and concepts that each student may be struggling with. Such systems also would give the same grade to students who half-understood all the concepts as to those that understood half the concepts really well and did not understand the others at all. In other words, implicitly or explicitly, the courses' lessons are treated as equally important. We did not believe this to be true for newcomers to principled negotiation. Some concepts and skills are more important to learn than others, and mastery of a few lessons can produce better outcomes than a broader, but less clear, understanding of many. The assessment system's effect on students' learning and priorities (see Ebner, Efron, and Kovach, *Evaluating our Evaluation*, in this volume) should be leveraged to reflect this and to encourage students to prioritize correctly. As students prioritize lessons, they can also target a grade and tailor their efforts to that grade. Instead of completing all of the assignments but with the minimal effort to get their target grade, students in the benchmark system need only complete those assignments necessary for the grade. This allows students to manage their efforts better across multiple courses in the semester, and saves the instructor from reading half-hearted assignments.

The course in which we implemented the benchmark system is a basic negotiation course taught over a whole semester. The number of students in this class is capped at thirty-six. The students are all post-graduate and the class typically includes people from all over the world, with the majority coming from Asia. The students primarily come from the School of Public Policy, though a few are students doing dual degrees at this school and another, or are exchange students from other universities. There are no prerequisites for participating in this course.

In the next section, we will describe the course and benchmark grading system experiment that we implemented in fall semester 2010. Then we will provide some data and observations on the outcomes of the experiment, and conclude with what we have learned from this experiment.

## The Benchmark System

The benchmark system we used established two types of mileposts that students had to pass. We termed the first kind of benchmarks “absolute benchmarks,” meaning that students could not achieve a certain grade unless they demonstrated convincingly that they understood one or more particular concepts or skills. The second kind we termed “accumulative benchmarks,” meaning that we added the number of skills or concepts the students acquired and used to determine their grade. The accumulative benchmarks were used for the higher grades (B+ to A) and allowed students some flexibility about which skills they chose to improve as they sought higher grades. Table 1 below provides a summary of the grades and what benchmarks students needed to pass in order to achieve them.

**Table 1: Grades and Benchmarks Requirements**

Grade	Requirements to pass the benchmark
A	All three benchmarks outlined in the B+ grade (1 – 3 below) <u>AND</u> Capable of negotiating beyond positions
A-	Two of the benchmarks outlined in the B+ grade (1–3 below) <u>AND</u> Can develop 2 or more possible packages making use of creative options and trades
B+	One of the following: 1. Full preparation memo <u>covering all elements well</u> (15 pages double-spaced; approx. 3750 words). <u>OR</u> 2. Reflective Learner (a good memo showing that they can learn from experience) (3-4 pages double-spaced; approx. 750-1000 words). <u>OR</u> 3. Can do a good four element preparation <sup>1</sup> (10-15 pages double spaced; approx. 2500-3750 words). <u>AND</u> Everything below

B	Can do a GOOD four element preparation (a grade of 4 out of 5 for <b>each and every</b> of the four elements on the quiz or your own analysis for a negotiation case). Can identify potential trades. <sup>2</sup>
B-	Shows that they understand what an interest is and how it is different from a position in practice. <sup>4</sup>
C+	Can generate creative options. <sup>4</sup>
C	Can give a standard definition of the seven elements.
C- D F	These are grades you can get if you pass none of the benchmarks above or the instructors drop your grade because you didn't show up for a negotiation simulation.

When we developed the benchmark system, we did so with several goals in mind:

- 1) **Transparency:** Our students often complain about the lack of transparency in how they are assessed, so we wanted to see if we could provide a grading system that clearly indicated: a) that there was a hierarchy and progression of concepts and skills; and b) what students would have to achieve to get certain grades.
- 2) **Hierarchy:** We wanted to establish a hierarchy of concepts and skills so that students had to learn the basics before they moved to more advanced material.
- 3) **Individual Monitoring:** Third, we wanted to know what concepts or skills each student was struggling with.
- 4) **Practice/Concept Differentiation:** To get the lower grades, the emphasis was on students' *understanding* of conceptual elements. The tipping point grade was a B, in which the *skill* of preparing for a negotiation using some of the seven elements was essential. For the higher grades, students had to demonstrate both a mastery of certain concepts and the ability to put them into practice during a negotiation.
- 5) **Student Prioritizing:** We wanted to provide the means by which students might make trade-offs in how they allocated work across courses. Another frequent complaint among students was that there was too much work across the semester's set of courses. Many said that they never had time to *think* about what they were learning in more depth, because they were constantly doing assignments. We wanted them to be able, should they desire, to set a target grade, and thus manage their workload in this course should they need to spend more time on others. We also saw potential benefits in this approach for us, since if students could focus on a fewer number of assignments to get the grade they

wanted, we would also have to read fewer assignments, and would not have to grade assignments that students put little effort into.<sup>3</sup> This effort had risks, of course, since students might have gotten the message that this course was less important. The chosen approach was to make that choice easier, since our past experience had shown that students generally put significant effort into other assignments.

### **Absolute and Accumulative Benchmarks**

With regard to the mix of absolute and accumulative benchmarks, the rationale was as follows. We felt that students must understand certain concepts and skills before they could get each of the grades below a B+ grade. It was for this reason that the benchmarks for B and below are “absolute.” The hierarchy among the concepts and skills is deliberate. For example, we wanted them to have a basic understanding of the *Seven Elements framework for negotiation* (Patton 2005)<sup>4</sup> and so that was the benchmark for a C grade. We made being able to *generate creative options* the benchmark for a C+ grade, and the basic practice of *discovering interests* in the beginning of a negotiation the benchmark for a B-. We chose this order because the process of discovering interests, when the other party is not initially cooperating, is more difficult than creating options. Similarly, we made *identifying possible trades* one of the two components of the B grade benchmark because we found that students can generate new ideas to meet identified interests more easily than they can identify the need for, and then actually make, *trades across interests*, rather than *concessions on issues*. The ability to identify trade-offs is at the core of integrative thinking and practice, and as a result needed to be included in the core or basic knowledge package.

Above the B grade, the system was more complicated. To get a B+, the students had to pass two benchmarks. The first was the successful completion of one of the written assignments. The second was demonstrating their ability to use the various seven elements during a simulated negotiation with the instructors. Here, we were mostly looking at a “basic” level of skill, where students could apply the skills when we were relatively compliant.<sup>5</sup>

To get an A-, the students had to complete one additional written assignment successfully. And finally, an A grade (the highest possible grade in the system) required that students finish all three written assignments successfully *and* that they demonstrate their ability to search for interests and develop possible solutions when we stuck to our positions, resisted their initial options, withheld key interests and otherwise did not follow ideal interest-based negotiation principles.<sup>6</sup>

## Mastery, not Percentage

The benchmarks could not be passed with a so-so performance. To achieve each benchmark, the students had to demonstrate clearly that they had mastered the concepts and skills assigned to that grade. For example, for the quiz benchmark, the students needed to show that they could use interests, alternatives, options, and objective criteria in preparing for a negotiation. If one of those elements was used incorrectly or inconclusively, then they would not pass the benchmark. As we realized that this was a high bar to pass, we sometimes provided students with multiple opportunities to pass the lower benchmarks. For example, students were given three chances to pass the quiz that was part of the B grade benchmark and two opportunities to do the simulated negotiation interview assessment.

## Lessons Learned

As noted above, the benchmark system aimed to: 1) improve transparency in grade assignment; 2) reinforce students appreciation for the *hierarchy* of concepts and skills; 3) make it possible for the students and the instructors to track more clearly their specific strengths and weaknesses; 4) differentiate between conceptual understanding and practical application; and 5) help students balance their workloads better according to their goals regarding what grades they wanted to achieve in this course and others during that particular semester. While the benchmark system was useful for tracking the specifics of student understanding, it was only partially successful in how it allowed students to prioritize their efforts within the course and balance their workload across their semester's courses. It also had some other unanticipated challenges as well. These drawbacks will need to be overcome in order for this system to be effective in regular negotiation classroom settings.

To start with, the system was successful in helping us adapt the course as we tracked how certain skills and understandings were distributed among the class and over time. This was especially useful as we tried to make sure that as many students as possible left the course with at least some understanding of some basic concepts and skills.<sup>7</sup> On the other hand, we had to grade early assignments soon after they were administered, so that we could assess which students had not passed the benchmark and what concepts or skills were lagging for that individual and across the whole class. This placed a greater demand on our time management.

Many students said that they appreciated the transparency of the system, but they also found it hard to understand at times. While we explained the system multiple times during the semester, in classroom, in small group and one-on-one conversations, students

were still confused by the end of the semester. As best we can tell through direct interactions with students, the reasons for this are twofold. First, the system was a bit complicated, especially when students sought a grade above a B. As described above, these grades required doing well both in their simulated negotiation interview and in at least one of the alternative written assignments. Additionally, we think the system itself was different enough from traditional assessment systems that it was not as easy to remember. Even students whom we were sure understood the system after a one-on-one conversation with us would be confused several weeks later.

We realized, during the semester, that the transparency had another drawback. Many educational institutions require a certain grade distribution amongst a class (for example, using a bell curve). In this class, we initially worried that students might get too many high grades, and then, as many struggled with some of the basic benchmarks, that too few would do so. Either way, such a transparent system gives the instructor less room to adjust grades according to the school's needs and the instructor's understanding of each student's capability from non-assessment observations.

The absolute nature of some of the benchmarks was a drawback for other reasons as well. Consider the quiz that we used as part of the B grade benchmark. We set that benchmark quite deliberately as the minimum level of comprehension and application we wanted students to leave the class with. However, we realized early on that if a student was unable to pass that benchmark during one of the quiz attempts (we gave the quiz three times during the semester), then they would no longer be able to get more than a B- for the rest of the semester.<sup>8</sup> We felt that it was unlikely that they would try very hard or even attend the rest of the classes, should that happen. While this may seem fair to these individual students, it posed two problems for the class. First, instructors always hope that students will learn as much as possible during their courses.<sup>9</sup> Second, their departure could make future simulations harder to manage, though we circumvented this possibility somewhat by reserving the right to drop grades for those who did not participate in simulations without a good reason.

This realization that students might know too much about their grade before the class finished thus increased our efforts to make sure that all students passed at least the quiz part of the B benchmark. This commitment became too high, as one of the complaints of the more successful students was that we spent too much time reinforcing the basic concepts and skills in class and too little on the ones that followed. While, in the end, all students passed the B benchmark, in post-course reflections we believe a few may not

have deserved that pass. So, the benchmark system's design made it possible that some students would know too much about their grade before the semester ended, and it made our attachment to their grade too strong and too immediate. Finally, the multiple opportunities we gave students obviously consumed more of our time.

## Conclusion

The benchmark assessment system was intended to increase transparency of grading, motivate students with visible goal posts for higher grades, and make more accessible information about each student's and the class' strengths and weaknesses across the skills and concepts. Based on the lessons learned, we recommend some possible modifications of the benchmark assessment system here.

Perhaps the most important point of reflection is the use of absolute vs. accumulative benchmarks, and their impact on what skills students learn and when students know their grade. Both kinds of benchmarks require that certain skills and concepts be mastered to a high level as the students seek to improve their grades. However, absolute benchmarks allow the instructor to set the hierarchy of skills to be mastered, while accumulative benchmarks allow the students to choose which ones they really have to master.

Accumulative benchmarks can still be set up so that certain lessons are prioritized. For example, the instructor can still assign grades from D to B+ by the number among the seven elements that the students have mastered, without saying which ones have to be mastered first. This still says that the seven elements have to be mastered before whatever other "higher" lessons the instructor assigns to the grades A- and A, but still allows the students to choose the order of lessons they seek mastery of.

If, on the other hand, the instructor wants to make use of the more strict hierarchy of lessons favored by absolute benchmarks, then the timing of assessment becomes quite crucial. Looking back, we could have given students the last B benchmark quiz (or other assignment) at the end of the semester. However, this would have provided little time for the students who pass this last quiz to attempt some of the higher benchmark assignments, and it is unlikely that many would have started the assignments for the higher benchmarks until they knew the outcome of the B grade benchmark assignment. Whether or not they deserve that chance would be up to the instructor.

## Notes

<sup>1</sup> By “four element” preparation, the instructor means that the students should be able to develop a correct preparation sheet for an upcoming negotiation that covers interests, alternatives, options, and objective criteria.

<sup>2</sup> For these benchmarks, the students had to demonstrate the skills in their simulations or the simulated negotiation interview assessment.

<sup>3</sup> Using a traditional assessment system, on the other hand, encourages students who are content with a lower mark to submit all the assignments, but with less effort placed on each. In other words, the instructor has to read all the assignments, including many that are disappointing.

<sup>4</sup> For a description of the seven elements (relationship, communication, interests, options, standards, alternatives, and commitments), see Patton (2005).

<sup>5</sup> The details and lessons learned from these “simulated negotiation” interview assessments are described in Fuller, *Interviews as an Assessment Tool*, in this volume.

<sup>6</sup> While this experiment was designed for teaching interest-based negotiation, we could see no reason why it could not be used as a means to assess other approaches to negotiation or any other course for that matter.

<sup>7</sup> To our surprise, for example, students had the greatest difficulty applying objective criteria when they prepared. We used that information to cover that topic again in the later classes.

<sup>8</sup> In the first round, only thirty-three percent passed, in the second forty percent passed, and the last ones passed on the final quiz.

<sup>9</sup> The use of assessment as an intentional lever for *encouraging* and guiding student learning is a tenet underlying many if the chapters of this book. The notion that an assessment method might actually *discourage* learning calls for further thinking and careful application.

## References

- Patton, B. 2005. Negotiation. In *Handbook of dispute resolution*, edited by M. L. Moffitt and R. C. Bordone. San Francisco: Jossey-Bass.

