TRACKS IN A MATH COURSE

FRANK QUINN

Abstract. Variation in student interest, preparation, and performance is usually accommodated by offering courses at several different levels and placing students in them at the beginning of the term. This practice has serious drawbacks that might be avoided by reversing the placement strategy. In a tracked course students enroll in a combined course, sort themselves into tracks according to performance, and the decision about the level they receive credit for is made at the end of the term. Resource constraints will make this approach impractical in many cases, but when it can be used it could significantly improve outcomes.

Contents

1. Introduction

Our students are not well served by traditional course design. They come to us with diverse backgrounds, interests, degrees of engagement, and ability, but courses are one-size-fits-all: uniform assignments, tests, and—in principle—uniform grading scales. Few students receive optimal instruction and a significant number are seriously out of place. Student/course mismatches shortchange students and reduce course effectiveness.

Date: October 2008.

The traditional way to address this problems is to offer several courses on essentially the same material but at different levels, say Standard and Advanced. However each course still has the problems on a lesser scale, and there are misplaced students: the best students in Standard should be in Advanced, and those who make F or D in Advanced should have been in Standard. We explain in §3 that this is an inherent problem with multi–level courses, and in particular cannot be addressed with better placement tests. There are also D and F students in Standard who should be in a lower level if one were available, and the best students in Advanced should be in a higher level. Offering multiple levels helps but doesn't solve the problem.

Our suggestion, in a nutshell, is to offer tracks at different levels *in a single class* rather than in separate classes. This would improve mobility between levels and avoid the misplacement problem. And if it can be made to work at all it should be possible to offer tracks at more levels than practical with separate courses.

The essay begins with discussions of the problems to be addressed. Performance diversity in a single class is discussed in §2 Outcome Diversity, and problems with resolving this through placement classes at different levels are described in §3 Placement Tests Are Not The Answer.

The main idea is described in §4 Tracked Courses, using perspective from §§ 2–3 on the problems to be avoided. As usual with a clever idea the real question is whether or not it is practical. Some of the many difficulties are described in §5 Implementation. It is doubtful that this could be implemented in a traditional class with current student/teacher ratios. There are, however, long–shot scenarios, and the benefits would be so great that these are worth considering.

2. Performance Diversity

The bottom line in a math course is end-of-course performance. For example, students who make grades of F or D were in some way not well matched to the course and probably should have been in a different level. Outcomes can't be used to identify these students at the beginning of the course, and the ramifications of this are discussed in the next section. Here we discuss problems resulting from having students who *will eventually fail* whether we can identify them ahead of time or not. We also describe problems of a very different nature at the other end of the spectrum.

2.1. **Under–performing students.** Roughly speaking, under–performing students are those who end up with grades of F or D in a course. Actual grades are not quite the right measure because grades are often adjusted to avoid having a lot of Fs and Ds. A better description is: students who would have gotten F or D in the absence of such adjustment.

The big problem associated with under–performing students is that changes made to accommodate them undercut learning of other students. We begin with under–performing students in Advanced courses.

Skill–oriented math classes are important because skills—the ability to work problems—are vital for success in later coursework, and eventually for use of mathematics in technical professions. Grades in a skills course are supposed to reflect acquisition of skills. If there are a lot of students who do not acquire skills then in principle there should be a lot of failing grades. Before the 1970s (roughly) this was standard practice and these courses often did have high failure rates.

High failure rates are now considered unacceptable. Many factors contribute to this but one is the realization that skills courses are neither appropriate nor necessary for many students. Under–performing students may be misplaced, for instance to "offer them an opportunity", not dumb. There are limits to how much it is reasonable to punish them for simply being in the wrong place or unable to take advantage of an opportunity. In any case a number of adjustments have been made to lower failure rates:

- expectations have been generally lowered;
- imprecise "understanding" may be accepted when skills are poor; and
- practices like extra credit, grade curves, dropping the lowest test, and soft homework scores are used to disconnect performance and grades.

The result is that course goals and grades have become ambiguous. Grades no longer indicate acquired skills, and it is unclear even to the best students that skills should be the key objectives. In effect the course goal has been changed to include "useful exposure", and while this may provide general life benefits it does not prepare students for advanced work.

Standard–level courses also have students who get F or D grades, and course goals are distorted by adjustments made to accommodate them. These distortions cause less long–term damage than in Advanced courses because skills and preparation are not the main objectives. Further, these students are usually not "misplaced" in the sense of being in the wrong course because there usually is no lower–level course. We do *not* want to interpret "misplaced" to mean "don't belong in any math course". This is a qualitatively different issue than being in the wrong course and it is not appropriate for us to address it here.

2.2. **Over–performing students.** Over–performing students are ones who would have been successful in a higher–level course. Since this concerns hypothetical outcomes in a course they didn't take, we can't identify specific students as over–performers. In particular a top grade in the course they did take does not reliably identify over–performers.

Problems associated with over–performing students differ depending on level, and differ from under–performing problems in that they do not effect course goals.

2.2.1. *Over–performing students in Standard courses.* Traditionally the main reservations about multi–level course offerings concern insufficient upward mobility: students who at some point got put in the Standard level and can't get out even though they would have done fine in Advanced. These students have, in a sense, been shortchanged by being deprived of opportunities available at the higher level.

This is an individual–benefit problem for specific students and does not have a negative impact on learning of other students.

2.2.2. *Over–performing students in Advanced courses.* These students are identified in even more hypothetical terms: success in some sort of "Gifted" course that generally isn't even offered.

The big problem in this area is societal. The huge role technology now plays in our lives means we need a significant number of extremely capable people trained to the full extent of their ability. Our educational system is not meeting this need. The few specialized high schools that do offer "Gifted" courses produce far more than their share of first–class scientists and engineers, even allowing for selective admissions. This suggests that the lack of very high–level courses in K-12 is a major part of the problem.

These students do not have a negative impact on other students. Further they have plenty of other opportunities so being shut out of the top echelons of science and technology is not a serious individual–benefit problem. It is only the societal problem that is severe.

2.3. **Summary.** Under–performing students impair the effectiveness of our educational system. Over–performing students are being denied opportunities, and in some cases being shut out of urgently needed technical leadership roles. Addressing these problems would seem to require offering instruction at more levels, and enabling easier and more appropriate mechanisms for mobility between levels.

3. Placement tests are not the answer

Most educators feel that if misplaced students are a problem then better placement tests are the solution. A key point is that placement has to be done at the beginning of the course, even if problems are most directly related to end-of-course performance. Justifications are:

- Placement decisions are limited by what we can measure. We can seek the best predictors of performance within this constraint, but there is no point in complaining about the limits imposed by this constraint.
- Since performance predictions are necessarily imprecise we should give students the benefit of the doubt. In other words deny admission to an Advanced course only if we are pretty sure the student will fail, and placement instruments need only be good enough for this.
- For whatever reason, educators put a lot of faith in placement tests.

All of these arguments are flawed. Placement is limited by beginning-of-course measures only if it has to take place at the beginning of the course. Mid–course placement is one of the advantages of tracked courses; we expand on this in §4. Giving students the benefit of the doubt maximizes individual opportunity but also leads to distortion of course goals and reduction in overall learning, as explained in the previous section. The last point concerns belief rather than an argument, and the objective in this section is to show this belief is unfounded.

3.1. **False positives and negatives.** Placement decisions can fail in two ways: false positives are students who get Advanced placement but turn out to be under–performers; false negatives are students denied Advanced placement but who would have been successful at that level.

If the placement system has more than 10% false positives then, as explained in §2.1, Advanced teachers have little choice but to weaken the link between credit and performance. Therefore a high false positive rate undercuts the skills orientation of the course. Since the course then no longer meets stated goals, credit for it gives misleading input for later placement decisions and raises false positive rates in later courses.

The usual way to keep the false positive rate low is to have higher requirements. But this inflates the false negatives and in practice makes skills courses unduly inaccessible. The extreme is a placement test that can only be passed by those who already know the material. The false positive rate is near zero and the course would go very well, but the false negative rate is near 100% and the course serves no educational function.

Real–life placement methods are too imprecise for there to be any satisfactory balance between false positives and negatives.

3.2. **Tests are Untested.** There is little solid data on effectiveness of placement tests because the self–fulfilling way they are used makes them almost impossible to evaluate. False positive rates tend to be masked by instructors' changing grading criteria to keep failure rates acceptably low. False negatives are practically impossible to assess and usually ignored.

There are a great many factors that effect performance but are not measured by tests: procrastination, short attention span, poor work habits, not to mention alcohol, drugs, and the emotional turmoil of youth. It should seem silly to even hope for a test with false positives under 10% and an acceptable false negative rate. Nonetheless many educators seem to take it as an unexamined article of faith.

3.3. **An Example, and Gateway tests.** We describe a real–life example. Our second-semester engineering calculus course has 25–30 sections each semester. Some years ago a series of brief computer–based "skills" tests were introduced to assess learning consistency across sections. The first of these measured entry skills and was essentially a placement test, though it was not used that way. Data from several thousand students showed an impressive correlation between scores and course outcomes. This probably could have been used to justify using the test for placement, but the statistics hid an asymmetry. Essentially all students who failed the course had failed the skills test but the converse did not hold; most who failed the skills test did fine in the course. The test had a low false positive rate but a very high false negative rate.

The story takes an interesting twist. These tests are multiple–try. Each individual test is different, students can get unlimited practice copies, and the proctored version can be taken multiple times with the best score counting. The data showed that students who initially failed but kept trying until they passed did almost as well in the course as those who got a perfect score the first try. This was taken to mean that entry skills are not immutable things that can only be measured and sorted, but somewhat malleable.

The test is now used as a skill–boosting "gateway". Students who sign up for the course *must* pass the skills test in the first week to stay in the course. Most pass on the first try, but:

- A few percent of enrollees drop out without attempting the test for credit. Presumably they have decided—after looking at practice tests—that they will not be successful, so the test is helping with *self*–placement!
- A tiny number, less than 1%, attempt the test but are unable to eventually pass.
- The remainder—the false negatives of the originial test—have to work to get their skills up to speed but do manage it.

Instead of a filter the test has become an instructional tool.

The tidy outcome in this example may depend on pre–filtering by the university admission process. Even so it does not reduce false negatives and positives enough to solve the basic problems of placement.

4. Tracked Courses

In a multi–level course students are sorted and placed in different classes for which they receive different credits. Tracked courses reverse this: students enter a single class, sort themselves into tracks as the course progresses, and only at the end of the course is a decision made about the credit received.

4.1. Basic Plan.

For simplicity we describe a course with two tracks: Upper and Lower.

- Students entering the course are not assigned to a track, and beginning classes are not specialized to either track.
- Tentative track assignments are based on performance on the first major test. Lower–track students who want to be in the Upper track can retake an equivalent test to try to get the necessary score. Students who qualify for the Upper track can, if they insist, be reclassified as Lower–track.
- Subsequent tests are track–specific. Upper–track students with unsatisfactory scores are reassigned to the Lower track, but again with an opportunity to improve the score.
- If there are several sections of the course then sections can specialize after the first test. Students might have to change sections to be in a class appropriate for their current track.

At the end of the course each student will be in one of the levels, and will have test and other course scores. Grading and credit is handled as follows:

- Upper–track students receive Upper–track credit that qualifies them for more–advanced later courses, and grades A, B, or C depending on scores. Students who would have gotten grades indicating unsatisfactory Upper–track performance have dropped to the Lower track where more appropriate standards can justify better grades.
- Lower–track students receive lower–track credit and the usual A–F grades, unless there is a yet lower level or track for the under–performers.
- There is an element of choice for Upper–track students: any Upper track grade can be converted to an A in the Lower track.

The choice offered in the last point provides a safety net. Students interested in law or medicine or seeking admission to elite college or graduate programs often avoid serious math courses to avoid damage to their grade point average. This is unfortunate because these students are often quite capable of Upper–track work. They might even be lured into a technical profession: many people in mathematics and science ended up there because they took a tough course and liked it.

An important feature of this design is that marginal students make their own decisions. Students struggling to stay in the Upper track may decide they aren't *that* interested in technical careers anyway, and change their goal to getting an A in the Lower track. This is certainly better than going limp and dragging down the whole class. If they are determined to stay in the Upper track they are motivated to work harder and rise to the right challenge. Finally this decision is made in small steps—one test at a time—so they can see exactly what is required and make an informed decision.

4.2. **Introduction to Proofs course.** The previous section is implicitly aimed at K–12 and the first few years of college. This section suggests that the approach could also be useful at higher levels.

In traditional college math sequences there is a shift of emphasis after calculus from problem–solving to more abstract and conceptual reasoning, “proofs” for short. Most students find the transition to proofs uncomfortable and by and large only math majors attempt it. This is unfortunate since the generalized reasoning skills acquired this way are germane to cutting–edge work in any science or engineering field. Some top software companies, for instance, recruit PhD mathematicians on the principle that it is easier to teach someone with high–level logical skills to use computers than it is to teach a computer expert to think on a higher level.

Until relatively recently the custom was to introduce students to proofs in a sink-or-swim way in courses on real or complex analysis or modern algebra. This was tough on students but satisfactory numbers made it through. Expansion of graduate programs in the 1970s and later weakening of lower–level education made this approach unworkable and many programs introduced “Introduction to Proofs” courses to help with the transition.

There is a new difficulty with the problem–proof transition. Further softening in lower–level courses has meant that there are fewer students with the preparation and discipline to make the transition, even among math majors, and even with help from a Proofs course. Faced with the need to keep numbers up and programs viable, some departments have softened their proofs courses. In effect they offer exposure credit to boost low skills scores. Naturally this degrades the end product.

Some university undergraduate math programs are now almost incapable of producing students that would qualify for their own graduate programs. Elite graduate programs sustain quality by recruiting foreign students. Many less–elite graduate programs are being softened to be accessible to Americans because the alternative is to close down. In other words the upper end of our mathematics educational system is starting to erode.

Using tracks, for instance Professional and General, in a proof course would help with this. The General track would be quite satisfactory for prospective K-12 teachers and the less math–intensive sciences. The Professional track would require the discipline needed for further math and math–intensive science and engineering, without harming the General–track students.

4.3. **Both Tracks and Levels.** We have used “multi–level” for separate courses with placement at the beginning of the term, and “tracked” for a combined course with placement at the end of the term. Up to this point the two have been compared directly in order to make the differences clear. In practice, however, the two approaches are complementary and often would be used together.

- A class with a serious skill component will spend a lot of time doing things non–skills students generally dislike. A lower track would reduce the impact of skills materials but would not make it more relevant. A non–skills course can focus more on interest and enrichment. Separate courses for the two levels are appropriate.
- There could be tracks in each level, with significance partly defined in terms of subsequent level placement. Upper–track credit in the upper level would

be required to qualify for the upper level in the next course; lower–track students would move to the lower level.

- This does re–introduce the mobility problem that is one of the virtues of tracks. It should be less problematic because students have had a lot of input into their placement, but some sort of upgrade process should be provided.

5. Implementation Problems

We list a few obstacles to implementation of tracked courses. Familiar problems such as developing texts and syllabi are not discussed.

5.1. **Resource Constraints.** Informal use of tracks was common in the one-room-schoolhouse days because there were too few students to justify separate classes. This is rarely possible now because it requires unrealistically low student/teacher ratios, willing and well–behaved students in K-12, and may require lower content density in college courses.

Formal introduction of tracks will not solve the student/teacher ratio problem. Teachers with typical–size classes are rarely able to focus on a subgroup for an extended time. Further, if test preparation, grading, etc. are done by hand then tracks could double the time required for this. Trying to introduce tracks in such cases will predictably lead to failure and should not be attempted.

Possible exceptions are:

- “Gifted” tracks in upper–level courses. The student/teacher ratios are generally low, students are cooperative, and very few students would be involved in the upper track.
- Lower tracks in low–level courses. These could be accomplished simply by changing the grading scale at the lower end, without changing materials or presentations.
- Computer–tested courses.

5.2. **Institutional Barriers.** Most institutions will be uncomfortable with the idea of students signing up for a course without knowing which course it is. They may also be uncomfortable with leaving course–credit decisions (i.e. end-of-course placement) in the hands of the faculty. Resistance by credit score–keepers (Registrars et al.) will make trials of the approach difficult.

6. Conclusion

Traditional course structure evolved to support a single goal, and the traditional single goal was good outcomes for a relatively small number of students. Education now has another, conflicting goal: modest outcomes for essentially everyone. The traditional structure has been unable to do justice to both goals at once. Tracks may provide a way to resolve this by offering several grading criteria and letting students play a significant role in deciding which goal and associated grading criterion is best for them.

Tracks should be relatively easy to implement in computer–tested courses. The extra burdens of course administration and multiple assessments make the approach infeasible in most traditionally–tested courses.