



The perils of prescribed grade distributions: What every medical educator should know

Kenneth D. Royal¹, Thomas R. Guskey²

¹Department of Clinical Sciences, North Carolina State University, Raleigh, North Carolina, USA. ²Department of Educational, School and Counseling Psychology, University of Kentucky, Lexington, Kentucky, USA

Address for correspondence:

Kenneth D. Royal,
Department of Clinical Sciences, North Carolina State University, Raleigh, North Carolina, USA.
E-mail: kdroyal2@ncsu.edu

Received: November 26, 2014

Accepted: March 24, 2015

Published: March 24, 2015

ABSTRACT

A common practice in medical education is to create a prescribed distribution of grades or ratings so that only a certain percentage of students receive the highest marks. This approach typically is employed to curb grade inflation and as a means to help faculty distinguish outstanding performers. Despite the well-intentioned reasoning for using prescribed grade distributions, a number of associated problems and probable consequences may result from this practice. Thus, the purpose of this article was to discuss the assumptions underlying this potentially unwise practice, the defensibility of this evaluation practice in the high-stakes arena of medical education, and the unintended consequences that such practices and policies may have on a student body and the larger educational environment.

KEY WORDS: Assessment, grading, medical education, medical education policy

INTRODUCTION

A common practice in medical education is to create prescribed grade distributions. That is, faculty, often at the request of a department chair or dean, attempt to sort students' grades into various levels of performance. Typically, only a certain percentage of students will receive the highest grades, another percentage of students will receive a less distinguishable grade, and so on. The practice is primarily intended to curb grade inflation; a problem that is particularly pervasive in clerkships [1], as many instructors fear providing too many high marks is inappropriate or would result in suspicion among others who may review course performance statistics. Proponents of prescribed grade distributions typically site the ability to distinguish outstanding performers as desirable and like the idea of having consistent grade distributions from year to year. Despite the perceived advantages of prescribed grade distributions, the purpose of this article was to discuss the assumptions, defensibility, and probable consequences of this practice.

PROBLEMS WITH PRESCRIBED GRADE DISTRIBUTIONS

Medical education has long been regarded as one of the most competitive disciplines to gain admission. With acceptance

rates generally around 5% or less, only the most talented, highly select, and highly motivated students typically are offered admission [2]. Given this backdrop, is it not illogical to anticipate that only, say 15%, of students will receive an "Honors" grade or otherwise "perform at the highest level" once admitted into a program? This is equivalent to saying 85% of academically elite students are expected to perform at some level below the "highest level" upon entering a program. This message tends to insinuate that either the program will do something to students that causes them to demonstrate a wider variation in abilities, or the program will otherwise make such judgments regardless of variation among students' abilities. What does this truly say about a program that chooses to adopt such a policy? What message does it convey to students?

When considering the well-known relationship between student learning and quality of instruction, there is even more confusion. Because performance measures are not only influenced by student ability, but also quality of instruction, then the assumption that only 15% of the students will "perform at the highest level" suggests there must be virtually no variation in instructional quality across courses. Thus, if the policy is applied across a department, or perhaps college-wide, then it would suggest each instructor is as gifted and effective as the next, and that each instructor's instructional techniques result in the

same proportion of students performing at the highest level. A review of course and instructor evaluations would assuredly dismiss this possibility.

If a program admits only the highest caliber students there will presumably be very little differentiation among students' performance unless instructors use assessment instruments that are both incredibly rigorous and evidenced to be capable of discriminating those students that learned more excellently than others. Given most medical educators have very limited, if any, training in psychometrics, it is difficult to imagine that most examinations conducted in any medical school possesses the type of targeted difficulty and psychometric qualities necessary to justify such grading distinctions. Without such validity evidence, a program or institution may have a difficult time defending their grading practices. What risks might this generate? Would this approach be legally defensible in such a high-stakes profession?

PROBABLE CONSEQUENCES ASSOCIATED WITH PRESCRIBED GRADE DISTRIBUTIONS

Prescribed grade distributions are largely incompatible with standards-based curricula as they primarily are based on norm-referenced assessment approaches. Norm-referenced designs are problematic for medical education for a number of reasons, not the least of which is unhealthy competition. Prescribed grade distributions make learning highly competitive for students, as each compete for the few scarce high grades awarded by their instructors. It conveys a message to students that performing well does not mean learning excellently; it means outperforming one's classmates. This type of competition can discourage student collaboration, make students reluctant to interact and learn from one another, and make instructors reluctant to provide individual assistance to students in fear that other students might misconstrue their actions as biasing the competition. Perhaps most devastating of all, unhealthy competition can also destroy the culture of a student body, as feelings of resentment, animosity, and jealousy may cause some students to behave inappropriately (e.g., rude and bullying behavior, sabotaging other students' work, withholding educational resources and materials, and so on). The effects of this competition model are so significant that long after physicians enter into medical practice many continue to inquire about their performance relative to others on medical certification examinations [3]. In fact, the American Board of Family Medicine has moved to no longer reporting percentile ranks, as the purpose of such examinations is not to measure one's performance relative to other examinees, but one's performance relative to a minimum passing standard [3].

CONCLUSION

Competition certainly has its place in the classroom. However, in a true and healthy learning environment, students should not be competing against other students for a few pre-determined number of high grades. Instead, students should be working with other students and competing against rigorous academic standards. Competition against academic standards provides an opportunity to unite students and their instructors with a common goal. This is far more likely to result in a positive academic environment, as helping a classmate would in no way diminish a student's chance of success for earning the highest grade. It is possible that this assistance could actually enhance each student's success [4].

If medical educators wish to distinguish the ablest learners, then providing challenging standards with psychometrically sound assessments is the best way to do that. With challenging standards and robust assessments, grades will become meaningful and defensible, and whatever percentage of students that happen to receive the highest grades will be those who have demonstrated performance at the highest level. This would alleviate any need to dedicate a limited number of grades to students based on some arbitrary threshold that is most likely indefensible and unhealthy for the culture of the program or institution. Thus, if instructors focus on providing excellent instruction to challenging academic standards, they can rightly spend their energies developing talent as opposed to wrongly trying to sort the talent.

REFERENCES

1. Fazio SB, Papp KK, Torre DM, Defer TM. Grade inflation in the internal medicine clerkship: A national study. *Teach Learn Med* 2013;25:71-6.
2. AAMC. U.S. Medical School Applications and Matriculants by School, State of Legal Residence, and Sex, 2013. Available from: <https://www.aamc.org/download/321442/data/2013factstable1.pdf>. [Last retrieved on 03 Nov 2014].
3. Royal KR, Puffer JC. Criterion-referenced examinations: Implications for the reporting and interpretation of examination results. *J Am Board Fam Med* 2013;26:225-6.
4. Hwee OH. Towards a grade-free system. *The Star Online*. Available from: <http://www.thestar.com.my/News/Education/2014/03/02/Towards-a-gradefree-system/>. [Last accessed on 03 Nov 2014].

© SAGEYA. This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

Source of Support: Nil, Conflict of Interest: None declared.