



EDITORIAL

## **Reflections on Change: Educational and Institutional Implications of “Regression Toward the Mean”**

Change is central to our professional work. As teachers our core task is helping learners change (for the better, presumably). If our institutions are to be appropriately responsive to evolving social conditions and learner needs, they must undergo continuing change. Yet, most of us quickly learn that meaningful, lasting change in people and their institutions doesn't happen automatically or easily. And constructive institutional changes often diminish over time.

Many educational institutions, in my view, retreat over time from excellence toward mediocrity. If we can agree that declines in our educational programs can happen, we must find ways to help prevent such declines. A first step is understanding the process. I propose that we need enlarged, continuing deliberations about these issues. This editorial is an effort to contribute to such a dialogue by focusing on one aspect of change: “regression toward the mean”. I invite your participation.

### **Implications of “Regression Toward the Mean” for Teaching**

Done properly, teaching is linked to student performance. We should select our instructional strategies from our observations of what students do. The way that works, and the implications of regression toward the mean (R2M) for the process, is well captured in the following anecdote, found on the internet (R.H. Williams, undated):

Not being a mathematician myself, I owe my understanding of “regression to the mean” to Kahneman and Tversky, who experienced an epiphany while trying to convince a group of Top-Gun pilot instructors of the intrinsic value of praise. When Kahneman & Tversky told the instructors that their students would respond faster to praise than to ridicule, the instructors were unanimously adamant in their ferocity against the advice. “We’ve tried that stuff and it doesn’t work!” they shouted. “Any time a student makes a perfect, three point landing and is praised for it, you can bet the next landing will be worse. But if the same student makes a horrible landing and you yell at him for it, the next one

will usually be better.” It was precisely then that the light bulb appeared over Tversky’s head. “Gentlemen, what you’re witnessing has nothing to do with your reinforcement, either positive or negative. What you’re seeing is merely regression to the mean. When a student makes a perfect landing, it’s nearly certain that the next one will be worse, regardless of whether the student is praised or punished. Likewise, a horrible landing will almost always be followed by a better one, no matter what you do. What we’re trying to teach you is how to more quickly raise the quality of the *average* landing.”

Teaching, at its essence, is a process of facilitating constructive change in the capabilities and performance of others. If we are to be optimally effective as teachers, we must understand the change process. And, as the anecdote above nicely illustrates, part of understanding change in learners is understanding the phenomenon of R2M.

## Understanding “Regression Toward the Mean”

Regression toward the mean is a statistical phenomenon, first identified by Sir Francis Galton (1886) from his studies of the heights of parents and their children. Whenever parents are well above or below average in height, their children are also likely to be similarly above or below average in height, but not by as much. In general, R2M predicts that more extreme (deviant) measures tend to be closer to the population’s average when a follow-up measurement is taken. Some examples:

- An athlete who has played exceptionally well in one season (or game) infrequently plays as well in the next.
- A rock group that has a highly successful initial CD is unlikely to do as well on their next recording.
- Students who do poorly on a test tend to finish closer to average on the next test.

In other words, the changes that take place over time among exceptional individuals and groups, and across generations tend to be toward the average for the larger population from which they are drawn. Some shifts occur because of the natural tendencies of biological organisms (as in Galton’s studies). Many shifts toward the mean occur because the measurements involved are composed of two parts: a valid part and a random error part. In much of educational testing, extreme results on one occasion are likely due in part to a random error component. By chance alone, that component is likely to contribute less on the next occasion.

## Implications of R2M for Testing and Research

The *true* component of a score reflects what the test is validly measuring and tends to be stable from one measurement to the next. The *error* component includes extraneous factors, such as the test-taker's emotional state at the time, lack of concentration, inappropriate testing conditions or improper question wording. This error component is the part that averages out from one measurement to another (Shaughnessy & Zechmeister, 1990). The less reliable a test, the more likely R2M will occur. And most classroom tests have had little if any systematic, psychometric development, so their reliability is often low.

Those of us who do research or test students in our courses must understand and account for R2M. If not, R2M can lead to misinterpretations of findings and incorrect conclusions. For example, students with the highest and lowest grades in a midterm exam tend to have scores closer to the mean on the final exam. This effect may be incorrectly interpreted as being a result of the program.

When conducting research, R2M may affect the internal validity of our experimental design. We could mistakenly conclude that a significant change is due to the treatment when it may be more due to R2M. R2M plays a crucial role in any research involving the measurement of change (Campbell & Kenny, 1999). We need to be particularly cautious about R2M whenever subjects are selected on the basis of extreme pretest scores, as when selecting patients with the most severe symptoms for evaluating a treatment. R2M can be expected whenever there is non-random sampling or assignment in experiments.

Although we have become better at accounting for R2M in some medical research by carefully randomizing our selection of subjects and by using control groups, problems still occur (Bland & Altman, 1994). Adjustments for R2M are not always possible when conducting educational and other behavioural science research or when teaching. Simply put, we need a well-developed sense of caution (and humility) when interpreting findings about change when administering educational tests, when conducting uncontrolled, non-random data gathering, and when teaching.

## Metaphorical Implications of R2M for Institutions

Those of us who are close observers of health professions education have witnessed many institutions retreat from earlier innovations. Such retreats are sufficiently common that some observers have borrowed the phrase, regression toward the mean, to characterize these events.

There is a substantial leap in moving from understanding change in the context of testing, research, and teaching to understanding change in educational institutions. The tendency of institutions to slip away from earlier instructional innovations is not the same as statistical R2M. But the process

looks the same, so using the same name seems apt, at least metaphorically. While R2M among individuals occurs because of measurement errors and biology, the causes of institutional retreats from instructional innovations are more psychological and sociological. Still, understanding R2M can offer some helpful insights into the organizational change process.

Among the many factors that drag institutions toward returning to being average (toward mediocrity?) are the following:

- Education is often a secondary (or lesser) priority among those who pass judgment upon institutions that are educationally innovative. Many of those people are not regular readers of literature that would inform them of emerging ideas and findings that might support the directions the “deviant” institution is trying to pursue.
- Almost inevitably, institutions grow and that growth typically brings the arrival of people who achieved their success at more “average” institutions, doing less innovative education. Typically, these newcomers are less supportive of the innovative programs than are those who developed them.
- External professional organizations and accreditation bodies tend to be dominated by people who spent their professional careers in “conventional” institutions and they are often defenders of the ways they did things rather than admirers of departures from the “tried and true.”
- Because of the factors above and others, sustaining innovations often requires a level of energy and dedication that is difficult to maintain over time, especially if the reward systems are tilted toward bench research and clinical practice, rather than teaching, as they often are.

Successive moves from extremes toward averages among educational institutions, as when a medical school gives up on assigning students to extended stays in rural communities, is certainly understandable, given the conditions summarized above. Can backsliding from earlier, unusual instructional programs toward more conventional approaches be avoided? Well, R2M in institutions, as with individuals, is preventable. Achieving prevention, however, is not easy. Prevention of institutional R2M requires, among other elements:

- a widespread understanding of the risk that it might happen;
- an equally widespread understanding of the rationale for the change that was undertaken in the first place;
- persistent vigilance for signs of slippage in commitment;
- a carefully orchestrated process for reinvigorating the general dedication to innovation, which requires strong leadership, fresh ideas, instructional efforts, and an appropriate reward system; and
- a search for new faculty who will reinforce and support, not diminish, what is being done, including an investment of time and effort educating new faculty regarding the underlying rationale for the approaches being used.

In other words, sustaining innovation can require at least as much effort and may require a different set of skills than were needed for initiating the original change. We need more people and institutions that are willing and able to dedicate resources and leadership to the challenging but vital effort of preventing their program's regression toward the mean.

This journal and its sponsoring organization have been trying to do their part. The founders of this journal emphasized the importance of change by making it the key word in its subtitle. As a journal, we've given our fair share of attention to the topic of change (e.g. Cox, 1999; Guilbert, 2001; Tiberius, 2001). The people who established The Network that sponsors this journal recognized that our health professions education institutions need to undergo continuing change as an inescapable part of remaining current and effective. They routinely emphasize this point at annual meetings and in Network publications. If these efforts are to have lasting success, we need widespread participation and ongoing dialogue. If you aren't already doing so, I encourage you to join this effort.

## Note

1. <[http://www.wizardofads.com/archive/regression\\_to\\_the\\_mean.htm](http://www.wizardofads.com/archive/regression_to_the_mean.htm)>.

Hilliard Jason, MD, EdD  
 Editor, *Education for Health*  
 <HJason@mac.com>

## References

- BLAND, J.M. & ALTMAN, D.G. (1994). Statistics notes: some examples of regression towards the mean. *British Medical Journal*, 309, 780.
- CAMPBELL, D.T. & KENNY, D.A. (1999). *A primer on regression artifacts*. New York: Guilford Press.
- COX, K. (1999). Persuading colleagues to change: fifteen lessons learned from more than 20 years of trying. *Education for Health: Change in Learning and Practice*, 12, 47–353.
- GALTON, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute*, 15, 246–263.
- GUILBERT, J.-J. (2001). Curriculum change and strategies, past and present: why is it taking so long? *Education for Health: Change in Learning and Practice*, 14, 367–372.
- SHAUGHNESSY, J.J. & ZECHMEISTER, E.B. (1990). *Research methods in psychology*. New York: McGraw–Hill.
- TIBERIUS, R.G. (2001). Meeting the challenge of a changing teaching environment: harmonize with the system or transform the teacher's perspective. *Education for Health: Change in Learning and Practice*, 14, 433–442.