On Teaching and Learning

Editors:
JANINE BEMPECHAT
DEAN K. WHITLA

On Teaching and Learning publishes articles and essays on aspects of pedagogical practice and on research that has implications for teaching.

Letters and suggestions for articles or essays should be addressed to the Editors, 317 Science Center, Harvard University, Cambridge, MA 02138.

Published by the Harvard-Danforth Center for Teaching and Learning.
James D. Wilkinson, Director.

The Journal of the Harvard-Danforth Center

On Teaching and Learning

1
Faculty Development Writ Large: A Decade of Working with Teachers to Help Them Improve Their Teaching
Dean K. Whitla

10
The "Muddiest Point in the Lecture" as a Feedback Device
Frederick Mosteller

22
How Students Learn: Part II
Robert Rosenthal

28
Teaching by the Case Method: One Teacher's Beginnings
Nona Lyons

36
College Teaching is a Funny Business
Ian W. Brown
How Students Learn: Part II

Robert Rosenthal

The topic that I discuss in this paper is one that I have been devoted to for thirty years, ever since I ruined the results of my doctoral dissertation at the University of California at Los Angeles. These results could only have been obtained if I had unintentionally communicated to the research subjects my expectation of how they should respond. This, basically, is the concept of the self-fulfilling prophecy — that one's expectations for another's behavior can be subtly communicated and ultimately realized. I have been in hot pursuit of this concept for these thirty years, trying to find not only the different contexts in which the self-fulfilling prophecy occurs, particularly in the classroom, but also how it operates. What is it that we do differently to students for whom we have high versus low expectations?

Some background on the concept of the self-fulfilling prophecy will help put my research into perspective. Gordon Allport, a Harvard psychologist, studied self-fulfilling prophecies at the international level. He believed that the self-fulfilling prophecy was one of the causes of war. In his view, one country might see the other arm, then escalate the arms race to defend itself, and so on, until a war might finally erupt.

At a national level, in the field of economics, the sociologist Robert Merton felt that the 1929 stock market crash had much to do with the self-fulfilling prophecy. If one were to see many people in line at a bank waiting to withdraw their money, then one would most likely feel it necessary to do the same while there is still money to be had. One then becomes a stimulus for others to join the crowd. Before long, the bank really does not have the money, whereas at the beginning of the day, it may have been as sound as any other.

In the area of medical research, the placebo effect is a very well-known phenomenon. Interestingly, the efficacy of a new drug tends to decrease over time. Early in the history of a new drug, we tend to hear optimistic reports of its effectiveness. Physicians tend to communicate their enthusiasm for the drug to their patients, whose condition promptly improves.

In time, double-blind research experiments are conducted. Typically, neither the patient nor the doctor knows whether he or she is receiving the placebo or the new drug. The results of double-blind experiments are usually much less favorable to new drugs than are the results of non-double-blind experiments. With this new information, physicians are likely to communicate much less enthusiasm to new patients with the same disorder. For example, they may state their belief that the drug works for some, but not all, individuals. Given this context, it makes sense that later patients may not profit quite as much as early patients did.

Henry Beecher, the chairman of the Anesthesiology Department at the Harvard Medical School and at Massachusetts General Hospital, actually found that he was virtually the only scholar in the world who could not get morphine to work. On experimentally induced threshold level pain, he found that the morphine in his lab was ineffective. When he examined his studies in relation to others, he found that other researchers were performing their studies single-blind — the patient did not know who was getting the morphine or the placebo, but the doctor knew. Beecher's experiments were always conducted double-blind. Under these conditions, the morphine was no more effective than the placebo for threshold level pain. It appears that for this kind of experimental situation, the expectation of the physician makes a critical difference in the efficacy of the drug.

Even surgery has been shown to have placebo effects. The following is not the kind of research that institutional review boards would permit today. There have been experiments reported by Beecher and others in which they have performed surgery on an experimental group, and "sham" surgery on a control group. In the latter case, the operation takes place, but nothing is actually removed. It appears that for many disorders, patient outcome is just as good whether the surgeon opens and closes or opens, removes, and closes. Medicine is wonderful.

In the history of experimental psychology, the most important case study of a self-fulfilling prophecy is that of Clever Hans. In Germany at the turn of the century, there was a horse named Hans, owned by a German mathematics instructor, Mr. von Osten. This horse could tell time, read, discriminate colors, and solve mathematics problems and problems of musical harmony. Furthermore, the owner did not have to be there to send off cues, and one did not have to believe in the horse in order for it to perform its seeming intellectual feats. A commission of thirteen eminent scholars, circus keepers, psychologists, psychiatrists, and veterinarians investigated the horse and found that it was not
fraudulent. This horse really did what it was reported to be able to do.

This challenging case led two scholars, Pfungst and Stumpf, to conduct a series of investigations. They discovered, for example, that if one gradually increased the distance between oneself and the horse, it was not as accurate as often as if one were closer. They also revealed that if one knew the answer to a question, so did the horse; however, if one did not know the answer, neither did the horse. Initially, it appeared that this phenomenon was due to extra-sensory perception — this horse was not smart, but could read minds.

Pfungst and Stumpf finally solved the case of Clever Hans by carefully examining the behavior of the questioner. When one would ask the horse a question, it would tap out the answer with its hoof. In asking the question, the individual tended to look the horse in the eye and then look down at its hoof, because that was where it would be tapping out the answer. Looking down at the horse's hoof turned out to be the signal to the horse to begin tapping — half the secret to being a smart horse. The rest of the secret is to signal to the horse as to when to stop tapping. Even those who did not "believe in the horse" would begin to make subtle motor adjustments just as it got to the right number of taps. For example, they might straighten up a little. Pfungst noted that even the dilation of a nostril was enough to cue this horse.

In retrospect, it should not have been too surprising that this kind of thing could go on. In 1927, Bertrand Russell had written that "animals studied by Americans rush about frantically with an incredible display of hustle and pep, whereas animals observed by Germans sit still and think and evolve their solutions out of their inner consciousness."

My colleagues and I replicated the Clever Hans phenomenon, working instead with rats. We conducted what are known as maze-bright/maze-dull experiments, but with a twist. At the University of California at Berkeley, scientists have inbred colonies of maze-bright and maze-dull rats over several generations. This is done by breeding the fastest male and female, and the slowest male and female. Over several generations one achieves a non-overlapping distribution, such that every rat in the fast group is "smarter" than every rat in the slow group, for that particular maze. We told our experimenters that we had maze-bright rats and maze-dull rats from Berkeley. In fact, the night before the experiment, we skulked about the laboratory and randomly hung signs on the cages — half identified the rats as maze-bright and half as maze-dull. After one day, the rats that were believed by their experimenters — their teachers — to be bright, actually performed better than those that were believed to be dull.

After replicating these results several times, we wondered if we could get the same kind of result with school children. I expressed the belief that we could in an article published over twenty years ago in the American Scientist. I subsequently heard from a school principal in California, Lenore Jacobson, who challenged me to test my hypothesis in her elementary school. We collaborated on a study we called the Pygmalion experiment. In May, we tested every child in the school on a test we called the "Harvard Test of Inflected Acquisition," which was actually an IQ test. We told the teachers that this test predicted who was going to bloom intellectually in the coming academic year. The following September, we let the teachers know which children in their incoming class had scored such that they were going to "bloom" this year.

In reality, we did not score the test that May. Instead, like the maze-bright/maze-dull rats, we randomly assigned children to the group "that would bloom." At the end of the school year, those children who had been designated as showing academic potential gained more IQ points than did the children in the control group. This was an expected result, because, by that time, much evidence had accumulated indicating that this Pygmalion effect was a robust phenomenon. However, we also got a result that was completely unexpected. Not all children showed the same IQ gain; some had gained more than others. For those who had been expected to bloom intellectually and for those in the fast and average tracks, the more they gained in IQ, the more favorably they were viewed by the teacher. Of course, there were some children who had not been destined to bloom and who were also in the slow track who happened to gain in IQ over the course of the year. The anomalous, and somewhat frightening, result was that the more these children gained in IQ, the more unfavorably they were viewed by their teachers, with respect to both intellectual and socio-emotional evaluations. It appeared, then, that there were hazards to unpredictable, unexpected intellectual development.

There have been many replications of the Pygmalion experiment, both with children and adults. For example, a study at the Air Force Academy showed that mathematics performance improved dramatically when the instructors, without factual basis, were led to expect better performance from some sections of their math classes. Another study showed that welders, mechanics, nurses, pressers, and assemblers all did better work when expected to by the foreperson who was in charge of their training. Athletes do more push-ups and sit-ups, and swim bet-
ter, if they are expected to do so by the coach. (These are all experimental results as well.)

Experimental studies have also revealed interesting sex differences. For example, it is generally known that boys do not learn to read as well as girls. Not all teachers, however, are aware of this. For teachers who do not know this, it is simply untrue. Indeed, we find that when we examine how boys and girls learn to read, we can predict the magnitude of the sex difference from knowing the belief of the teacher. If the teacher knows that girls are better at learning to read than boys, then, for that teacher, the girls' performance will be better. But, if the teacher does not know this or does not believe it, the boys will perform as well as the girls.

What happens in classrooms that makes some children achieve more when they are expected to do so? A review of the literature has shown that four factors are involved. The first is climate. Teachers treat children for whom they have favorable expectations more warmly than they treat other children.

The second important factor is input. Teachers teach more material to children for whom they have higher expectations. We can argue that the long reading lists that we give out at Harvard can be seen as something very favorable to the intellectual development of the students. In making such demands, it is obvious that we are expecting great things of them. (Of course, there is probably a curvilinear relationship, such that if we assign too much reading, they will not do it.)

The third factor is output. The fact is that students from whom more is expected are called on more often and are given more opportunity to be responded to socratically. Mary Budd Rowe, an educational experimenter, has done research on response opportunity. She has found that the median length of time a teacher waits for a child to respond to a question is one second. She advised teachers to wait longer, and found that when they did so, they got higher quality responses. Not exactly a miracle.

The last factor is feedback. Differentiated feedback is found to be different for children for whom teachers have high versus low expectations. When teachers have high expectations, they let children know when they are wrong. When teachers have low expectations for children, they tend not to correct a bad performance. In a sense, it is a put-down to students when teachers do not insist that they do better.

The Pygmalion phenomenon occurs through nonverbal as well as verbal channels. We have increasingly discovered the importance of nonverbal behavior in classrooms, courtrooms, clinics, and experimental laboratories. Most recently, my colleagues and I conducted courtroom research in which we found that we could "postdict" the criminal history of the defendant from the tone of voice in which the judge spoke to the jury about the criminal. Of course, the jury is not supposed to know what the criminal history of the defendant is, but it was communicated through nonverbal channels. Similarly, our attitudes and feelings about our students are leaked to them through nonverbal channels of communication.

By way of summary, I quote from a teutonic predecessor of Pfungst and of Stumpf, Johann Wolfgang von Goethe who said, "If you treat a person as if they are what they ought to be, they will become what they ought to be and could be."
Parting Thoughts

I have touched on but a small part of teaching here. The mechanics as to what to do and what not to do are obviously based on personal experiences and are geared to my personality. I do feel, however, that knowing the options may be of some value to you in your work. The thing which I have failed to communicate is how little of all of this is actually work. As I look back through this paper, I am saddened to discover that I have not really revealed the joy of college teaching. I was trained to do research, trained to increase knowledge, but it means so little if I do not have the opportunity to communicate it. How often I have wanted to shake a student by the shoulders and blurt out the answer for what appears to me to be so obvious. That is the temptation of the teacher. But how much more satisfaction I have received when a student has slowly plodded through the reasoning process and discovered the solution for him or herself. Teaching: is indeed a funny business. How often you seem to be wasting time when you could just as easily do the job yourself. But as the years pass by and you see those very same students accomplishing things that you never dreamed of doing, you know you must be in the right business.

Acknowledgments

A more informal version of this essay was presented at the 1987 Fall Teaching Orientation of the Harvard-Danforth Center for Teaching and Learning. I would like to thank Eric Kristensen for inviting me to participate in this session and, thus, providing the stimulus for committing to paper what has been in my head.

The Contributors

Ian W. Brown is Associate Curator of North American Collections in the Peabody Museum of Archaeology and Ethnology. Mr. Brown teaches North American pre-history and a seminar on Indians in Museums at Harvard College. His research is on the archaeology of the lower Mississippi valley.

Nona Lyons is Lecturer on Education and Research Associate at the Harvard Graduate School of Education. A Spencer Fellow, Ms. Lyons teaches a course on teaching English, Social Studies, and the Humanities in secondary schools.

Frederick Mosteller is Professor of Mathematical Statistics, Emeritus at Harvard University. A gifted teacher, Mr. Mosteller has trained a generation of statisticians and has had an enormous impact on the study of statistics.

Robert Rosenthal is Professor of Social Psychology at Harvard University. A gifted teacher, Mr. Rosenthal teaches courses on research methods, statistics, and social psychology.

Dean K. Whitla is Director of the Office of Instructional Research and Evaluation in the Harvard Faculty of Arts and Sciences, and is Lecturer on Education in the Harvard Graduate School of Education. Mr. Whitla founded and directed the Harvard-Danforth Center for Teaching and Learning for the first decade of its existence.