WHAT YOU EXPECT IS WHAT YOU GET

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinates of pupils' IQ gains. *Psychological Reports, 19*, 115-118.

We are all familiar with the idea of the self-fulfilling prophecy. One way of describing this concept is to say that if we expect something to happen in a certain way, our expectation will tend to make it so. Whether self-fulfilling prophecies really do occur in a predictable way in everyday life is open to scientific study, but psychological research has demonstrated that in some areas they are a reality.

The question of the self-fulfilling prophecy in scientific research was first brought to the attention of psychologists in 1911 in the famous case of "Clever Hans," the horse of Mr. von Osten (Pfungst, 1911). Clever Hans was a horse that was famous for being able to read, spell, and solve math problems by stomping out answers with his front hoof. Naturally there were many skeptics, but when Hans's abilities were tested by a committee of experts, they were found to be genuinely performed without prompting from Mr. von Osten. But how could any horse (except possibly for Mr. Ed!) possess such a degree of human intelligence? Well, a psychologist, O. Pfungst, performed a series of careful experiments and found that Hans was receiving subtle unintentional cues from his questioners. For example, after asking a question, people would look down at the horse's hoof for the answer. As the horse approached the correct number of hoofbeats, the questioners would raise their eyes or head very slightly in anticipation of the horse completing his answer. The horse had been conditioned to use these subtle movements from the observers as signs to stop stomping, and this usually resulted in the correct answer to the question.

So, you might ask, how is a trick horse related to psychological research? Well, the Clever Hans findings pointed out the possibility that observers often have specific expectations or biases that may cause them to send covert and unintentional signals to a subject being studied. These signals, then, may cause the subject to respond in ways that are consistent with the observers' bias and, consequently, confirm their expectations. What all this finally boils down to is that an experimenter may think a certain behavior results from his or her scientific treatment of one subject or one group of subjects compared with another. Actually the behavior may result from nothing more than the experimenter's own biased expectations. If this occurs, it renders the experiment invalid. This threat to the validity of a psychological experiment is called the *experimenter expectancy effect*.

Robert Rosenthal, a leading researcher on this methodological issue, has demonstrated the experimenter expectancy effect in laboratory psychological experiments. In one study (Rosenthal & Fode, 1963), psychology students in a learning and conditioning course unknowingly became subjects themselves. Some of the students were told they would be working with rats that had been specially bred for high intelligence, as measured by their ability to learn mazes quickly. The rest of the students were told that they would be working with rats bred for dullness in learning mazes. The students then proceeded to condition their rats to perform various skills, including maze learning. The students who had been assigned the maze-bright rats reported significantly faster learning times than those reported by the students with the maze-dull rats. In reality, the rats given to the students were standard lab rats and were randomly assigned. These students were not cheating or purposefully slanting their results. The influences they exerted on their animals were apparently unintentional and unconscious.

As a result of such research, the threat of experimenter expectancies to scientific research has been well established. Properly trained researchers, using careful procedures (such as the double-blind method, in which the experimenters who come in contact with the subjects are unaware of the hypotheses of the study) are usually able to avoid most of these expectancy effects.

Beyond this, however, Rosenthal was concerned about how such biases and expectancies might occur outside the laboratory, such as in school classrooms. Because teachers in public schools may not have the opportunity to learn about the dangers of expectancies, how great an influence might this tendency have on the students' potential performance? After all, teachers have historically been given students' IQ scores beginning in the first grade. Could this information set up biased expectancies in the teachers' minds and cause them to unintentionally treat "bright" students (as judged by high IQ scores) differently from those seen as less bright? And if so, is this fair? Those questions formed the basis of Rosenthal and Jacobson's study.

THEORETICAL PROPOSITIONS

Rosenthal labeled this expectancy effect, as it occurs in natural interpersonal settings outside the laboratory, the *Pygmalion effect*. In the Greek myth, a sculptor (Pygmalion) falls in love with his sculpted creation of a woman. Most people are more familiar with the modern Shaw play *Pygmalion* (*My Fair Lady* is the musical version) about

the blossoming of Eliza Doolittle because of the teaching, encouragement, and *expectations* of Henry Higgins. Rosenthal suspected that when an elementary school teacher is provided with information (such as IQ scores) that creates certain expectancies about students' potential, whether strong or weak, the teacher might unknowingly behave in ways that subtly encourage or facilitate the performance of the students seen as more likely to succeed. This, in turn, would create the self-fulfilling prophecy of actually causing those students to excel, perhaps at the expense of the students for whom lower expectations exist. In order to test these theoretical propositions, Rosenthal and his colleague Jacobson obtained the assistance of an elementary school (called Oak School) in a predominantly lower middle-class neighborhood in a large town.

METHOD

With the cooperation of the Oak School administration, all the students in grades one through six were given an IQ test (called the *Tests of General Ability*, or TOGA) near the beginning of the academic year. This test was chosen because it was a nonverbal test for which a student's score did not depend primarily upon school-learned skills of reading, writing, and arithmetic. Also, it was a test with which the teachers in Oak School probably would not be familiar. The teachers were told that the students were being given the Harvard Test of Inflected Acquisition. Such deception was important in this case in order for expectancies to be created in the minds of the teachers, a necessary ingredient for the experiment to be successful. It was further explained to the teachers that the Harvard Test was designed to serve as a predictor of academic *blooming* or *spurting*. In other words, teachers believed that students who scored high on the test were ready to enter a period of increased learning abilities within the next year. This predictive ability of the test was also, in fact, not true.

At Oak School, there were three classes at each of the six grade levels. All of the 18 teachers (16 women, 2 men) for these classes were given a list of names of students in their classes who had scored in the top 20% on the Harvard Test and were therefore identified as potential academic bloomers during the academic year. But here's the key to this study: The children on the teachers' top 10 lists had been assigned to this experimental condition purely at random. The only difference between these children and the others (the controls) was that they had been identified to their teachers as the ones who would show unusual intellectual gains.

Near the end of the school year, all children at the school were measured again with the same test (the TOGA), and the degree of change in IQ was calculated for each child. The differences in IQ changes between the experimental group and the controls could then be examined to see if the expectancy effect had been created in a real-world setting.

RESULTS

Figure 1 summarizes the results of the comparisons of the IQ increases for the experimental versus the control groups. For the entire school, the children for whom the teachers had expected greater intellectual growth averaged significantly greater improvement than did the control children (12.2 and 8.2 points, respectively). However, if you examine Figure 1, it is clear that this difference was accounted for by the huge differences in grades one and two. Possible reasons for this are discussed shortly. Rosenthal and Jacobson offered another useful and revealing way to organize the data for these first and second graders. Figure 2 illustrates the percentage of the children in each group who obtained increases in IQ of at least 10, 20, or 30 points.

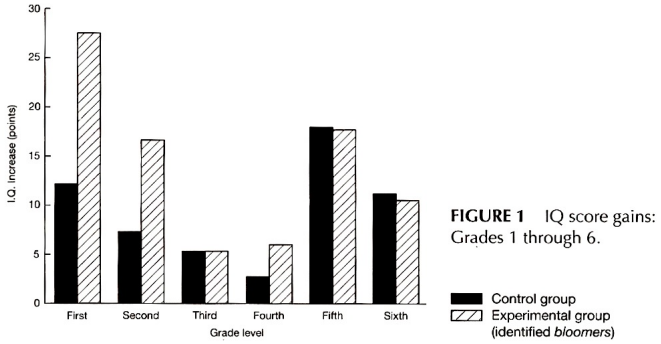

FIGURE 1 IQ score gains: Grades 1 through 6.

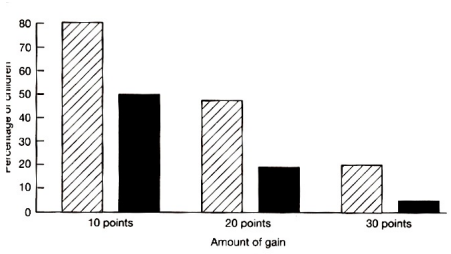

FIGURE 2 Percentage of first- and second-grade students with major gains in IQ scores.

Two major findings emerged from this early study. First, the expectancy effect previously demonstrated in formal laboratory settings also appears to function in less experimental, real-world situations. Second, the effect was very strong in the early grades, yet almost nonexistent for the older children. What does all this mean?

DISCUSSION

As Rosenthal suspected from his past research, the teachers' expectations of their students' behavior became a self-fulfilling prophecy. "When teachers expected that certain children would show greater intellectual development, those children did show greater intellectual development" (Rosenthal & Jacobson, 1968, p. 85). Remember that the data reported are averages of three classes and three teachers for each grade level. It is difficult to think of other explanations for the differences in IQ gains besides the teachers' expectations.

However, Rosenthal felt it was important to try to explain *why* the self-fulfilling prophecy was not demonstrated in the higher grade levels. Both in the article that is the focus of this chapter and in later writings, Rosenthal and Jacobson offered several possible reasons for this finding:

1. Younger children are generally thought of as being more malleable or "transformable." If this is true, then the younger children in the study may have experienced greater change simply because they were easier to change than the older children were. Related to this is the possibility that even if younger children are not more malleable, teachers may have believed that they were. This belief alone may have been enough to create differential treatment and produce the results that were reported.
2. Younger students in an elementary school tend to have less well-established reputations. In other words, if the teachers had not yet had a chance to form an opinion of a child's abilities, the expectancies created by the researchers would have carried more weight.
3. Younger children may be more easily influenced by and more susceptible to the subtle and unintentional processes that teachers use to communicate performance expectations to them:

> Under this interpretation, it is possible that teachers react to children of all grade levels in the same way if they believe them to be capable of intellectual gain. But perhaps it is only the younger children whose performance is affected by the special things the teacher says to them; the special ways in which she says them; the way she looks, postures, and touches the children from whom she expects greater intellectual growth. (Rosenthal & Jacobson, 1968, p. 83)

4. Teachers of these lower grades may differ from upper-grade teachers in ways that produce greater communication of their expectations to the children. Rosenthal and Jacobson did not speculate as to exactly what these differences might be if indeed they exist.

SIGNIFICANCE OF FINDINGS AND SUBSEQUENT RESEARCH

The real importance of Rosenthal and Jacobson's findings at Oak School relates to the potential long-lasting effects of teachers' expectations on the scholastic performance of students. This, in turn, feeds directly into one of the most controversial topics in psychology and education today: the question of the fairness of IQ tests. We'll return to this discussion shortly, but first, it is of interest to explore some later research that examined the ways in which teachers unconsciously communicate their higher expectations to the students whom they believe possess greater potential.

A study conducted by Chaiken, Sigler, and Derlega (1974) involved videotaping teacher-student interactions in a classroom situation in which the teachers had been informed that certain children were extremely bright (these *bright* students had been chosen at random from all the students in the class). Careful

examination of the videos indicated that teachers favored the identified *brighter* students in many subtle ways. They smiled at these students more often, made more eye contact, and had more favorable reactions to these students' comments in class. These researchers go on to report that students for whom these high expectations exist are more likely to enjoy school, receive more constructive comments from teachers on their mistakes, and work harder to try to improve. V\'hat this and other studies indicate is that teacher expectancies, while their influence is not the only determinant of a child's performance in school, can affect more than just IQ scores.

Imagine for a moment that you are an elementary school teacher with a class of 20 students. On the first day of class, you receive a class roster on which is printed the IQ scores for all of your students. You notice that five of your pupils have IQ scores over 145, well into the genius range. Do you think that your treatment and expectations of those children during the school year would be the same as your other students? What about your expectations of those students compared with another five students with IQ scores in the low to normal range? If you answered that your treatment and expectations would be the same, I'd be willing to bet that you'd be wrong. As a matter of fact, they probably *ɜhouldn't* be the same! The point is that if your expectations became self-fulfilling prophecies, that might be unfair to some of the students. Now consider another, more crucial point. Suppose the IQ scores you received on your class roster were *wrong*. If these erroneous scores created expectations that benefited some students over others, it would clearly be unfair and probably unethical. This is one of the major issues fueling the IQ controversy that rages today.

For many years, many researchers have charged that the standard IQ tests used to assess the intelligence of children contain a racial or cultural bias. The argument is that since the tests were designed primarily by white, upper middle-class males, they contain ideas and information to which other ethnic groups are not exposed. Children from various minority groups in the United States traditionally score lower on these tests than white children do. Since it would be ridiculous to assume that these nonwhite children possess less basic intelligence than white children, the reason for these differences in scores must lie in the tests themselves. Traditionally, however, teachers in grades K through 12 were given this IQ information on all their students. If you stop and think about this fact in relation to the research by Rosenthal and Jacobson, you'll see what a potentially dangerous situation may have been created. Besides the fact that children have been categorized in school according to their IQ scores (advanced placement, remedial classes, etc.), teachers' unintended expectations, based on this possibly biased information, may have been creating unfair self-fulfilling prophecies. The arguments supporting this idea are convincing enough that most states have instituted a moratorium on IQ testing and the use of IQ scores until such testing can be shown to be bias-free. And at the core of these arguments has been the research addressed in this chapter (see also the related work by Howard Gardner on *multiple intelligenceɜ*).

RECENT APPLICATIONS

Due in large part to Rosenthal and Jacobson's research, the power of teachers' expectations on students' performance has become an integral part of our understanding of the educational process. Furthermore, Rosenthal's theory of interpersonal expectancies has exerted its influence in numerous areas other than education. In 2002, Rosenthal himself reviewed the literature on expectancy effects using meta-analysis techniques (explained in the Smith and Glass study). He demonstrated how "the expectations of psychological researchers, classroom teachers, judges in the courtroom, business executives, and health care providers can unintentionally affect the responses of their research participants, pupils, jurors, employees, and patients" (Rosenthal, 2002, p. 839).

An uncomfortably revealing article incorporating Rosenthal's expectancy research examined the criteria school teachers use to refer their students to school psychologists for assessment and counseling (Andrews, et al, 1997). The researchers found that teachers referred African American children for developmental handicap assessment at rates significantly higher than the rates of Caucasian students in their classrooms. In addition, boys were referred in equally disproportionate numbers over girls for problems of classroom and playground behavior problems. The researchers suggested that the differences among the various student groups may have revealed more about teachers' expectancies than real individual differences.

Finally, Rosenthal's Pygmalion studies have not been without critics. Richard Snow at Stanford University has questioned Rosenthal's findings for over 30 years and the debate between them continues today (Rosenthal is at The University of California). A concise, pithy, and rather rancorous dialogue between them on this very topic appeared in a 1994 issue of *Current Directionɜ in Psychological Science*, the journal of the American Psychological Society (Rosenthal, 1994; Snow, 1994). It's a revealing and enjoyable read!

Andrews, T., Wisniewski, J., & Mulick, J. (1997). Variables influencing teachers' decisions to refer children for school psychological assessment services. *Psychology in Schools, 34*(3), 239-244.

Chaiken, A., Sigler, E., & Derlega, V. (1974). Nonverbal mediators of teacher expectancy effects. *Journal of Personality and Social Psychology, 30*, 144-149.

Pfungst, O. (1911). *Clever Hans (the horse of Mr. von Osten): A contribution to experimental, animal, and human psychology*. New York: Holt, Rinehart and Winston.

Rosenthal, R (2002). Covert communication in classrooms, clinics, courtrooms, and cubicles.
American Psychologist, 57, 839-849.

Rosenthal, R (1994). Critiquing Pygmalion: A 25-year perspective. *Current Directions in Psychological Science, 4*(6), 171-172.

Rosenthal, R, & Fode, K. (1963). The effect of experimenter bias on the performance of the albino rat. Behavioral *Science, 8*, 183-189.

Rosenthal, R, & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and pupils' intellectual development*. New York: Holt, Rinehart and Winston.

Snow, R. (1994). Pygmalion and intelligence? *Current Directions in Psychological Science, 4*(6), 169-171.