DOCUMENT RESUME

ED 023 774 UD 007 615

By -Gumpert, Peter. Gumpert, Carol

The Teacher As Pygmalion Comments on the Psychology of Expectation.

Pub Date Sep 68
Note -5p.
Journal Cit -The Urban Review, v3 n1 p21 -25 Sep 1968
EDRS Price MF -$0.25 HC -$0.35
Descriptors -Classroom Environment, *Disadvantaged Youth, Educational Experiments, Elementary School Students, *Intellectual Development, Perception, Psychological Characteristics, *Research Reviews (Publications), Student Attitudes, Student Teacher Relationship, *Teacher Attitudes, *Teacher Behavior

Although this review of "Pygmalion in the Classroom," an experimental study of the effect of positive teacher expectations on the intellectual development of their disadvantaged students, generally affirms the findings of the experiment, it points out that (1) since the average improvement of the experimental elementary school children was strongly associated with the improvement of the control children in the same classroom, the unit of analysis should have been the average intellectual gain of the children in the classroom as a whole rather than the gain of the individual child, and (2) in presenting their findings the experimenters suggested only the probable stability of their results but did not account for the apparent magnitude of their variables in the total context of their research. The review also speculates about how the teachers fulfilled the prophecy that positive expectations would increase student intellectual gains, particularly how they may have interpreted ambiguous events to confirm their positive attitudes and how increases in their interpersonal warmth may have lead to superior student performance. (EF)

The URBAN REVIEW Vol. 3 NO. 1 September 1968 | 21

The Teacher as Pygmalion: Comments on the Psychology of Expectation

by Peter and Carol Gumpert

In *Pygmalion in the Classroom,* Robert Rosenthal and Lenore Jacobson present an experimental investigation of an interesting and provocative question: do teachers' expectations of how well, or badly, their students will perform actually influence the intellectual development of the students? The background of the study is introduced by a reminder that prophecies are sometimes instrumental in bringing about their own fulfillment. This reminder is followed by a discussion of evidence for the operation of self-fulfilling prophecies drawn from medical, psychiatric, and behavioral science research. The evidence as compiled by Rosenthal and Jacobson is impressive: it is clear that what one person expects of another often has subtle, yet powerful, effects on the other. Since the teacher often "knows" (expects) which of her pupils will be brighter than the others, e.g., that white middle-class children will do better than the lower-class or Negro children, the critical question arises: can her expectation, *by itself,* bring about its own fulfillment?

To test the hypothesis that teacher expectations do, in fact, have such an effect, the investigators administered a group test of intelligence to several classes of children who were to enter grades one through six in 'Oak School.' The children's teachers were told that the test had been developed to predict forthcoming 'spurts' in intellectual growth; at the beginning of the fall term each teacher was given the names of several children in her classroom who could be expected to make unusual strides in intellectual development during the coming year. The specially designated children were actually no different from their contemporaries, having been selected randomly from the population of pupils at the school. The intelligence test was readministered to the children after a semester, after a year, and again after two years.

The major finding of the study was that the undesignated children, the control group, gained an average of more than eight IQ points after a year, while the special children, the experimental group, gained an average of more than 12. Statistical tests suggested that this difference was one that could not reasonably be ascribed to chance variation alone. Rosenthal and Jacobson concluded—and we concur—that their major hypothesis was con-

The Bettmann Archive

Rosenthal and Jacobson respond to this by postulating that their explanation for the Oak School result is not applicable to the earlier experiment, presumably because the schools in that study drew their pupils from a 'substantial middle-class community,' while Oak School's pupils came from a lower-class community. They do not explain how the socioeconomic

about its own fulfillment?

To test the hypothesis that teacher expectations do, in fact, have such an effect, the investigators administered a group test of intelligence to several classes of children who were to enter grades one through six in 'Oak School.' The children's teachers were told that the test had been developed to predict forthcoming 'spurts' in intellectual growth; at the beginning of the fall term each teacher was given the names of several children in her classroom who could be expected to make unusual strides in intellectual development during the coming year. The specially designated children were actually no different from their contemporaries, having been selected randomly from the population of pupils at the school. The intelligence test was readministered to the children after a semester, after a year, and again after two years.

The major finding of the study was that the undesignated children, the control group, gained an average of more than eight IQ points after a year, while the special children, the experimental group, gained an average of more than 12. Statistical tests suggested that this difference was one that could not reasonably be ascribed to chance variation alone. Rosenthal and Jacobson concluded—and we concur—that their major hypothesis was confirmed.

They also report a great deal of additional data on the more specific effects of the expectations they induced among the teachers of Oak School. Some of these ancillary results are intriguing, particularly some of the observations concerned with teachers' perceptions of their students during the first year of the experiment. However, looking carefully at all the results reported by Rosenthal and Jacobson—how the strength of teacher expectation varied with the sex of the students, their group status (minority or otherwise), their ability track, and so on—is a complex and confusing business. It is fairly easy to interpret each result taken alone, but the reader who attempts to make his interpretations consistent with one another will be in for a good deal of difficulty. The situation becomes even more confusing when we add to the Oak School results those obtained in the three similar studies in other schools reported by Rosenthal and Jacobson (pp. 96, 138, 145). Here we find apparent reversals of some of the effects that are given much emphasis in the Oak School study. One of these, to which Rosenthal and Jacobson pay special attention, is the result that after one year, Oak School girls in the experimental group showed greater gains on the reasoning subtest of the IQ test than did boys, while the reverse was the case on the verbal subtest, where boys showed evidence of greater improvement than girls. Since initially boys had shown slightly higher verbal scores and girls had shown slightly higher reasoning scores, Rosenthal and Jacobson interpret their results as indicating that each group tends to benefit most from favorable teacher expectations in those areas in which it is somewhat advantaged at the outset. However, in a previous experiment, while the initial scores for the sexes show the same pattern, the pattern of gains is reversed.

The Bettmann Archive

Rosenthal and Jacobson respond to this by postulating that their explanation for the Oak School result is not applicable to the earlier experiment, presumably because the schools in that study drew their pupils from a 'substantial middle-class community,' while Oak School's pupils came from a lower-class community. They do not explain how the socioeconomic class difference could account for the discrepancy in results between the two studies. The authors do not abandon their faith in the relationship between sex and subtest in either of the two studies, even though the results are by most people's standards mutually contradictory. They see the discrepancy in results as another demonstration of the "probable complexity of the operation of teacher expectancy effects" (p. 162).

This points to what we consider to be a major flaw of the book. The authors somehow give the impression that the 'expectancy effect' is a mysterious and profoundly complex entity, defying attempts by behavioral science to account for it. From such a viewpoint, it is not surprising that the importance of statistical significance is overemphasized by these investigators while the importance of replicability is underemphasized. We would have preferred their showing less tolerance for ambiguity and logical contradiction while pressing harder to find some consistent principles to explain their data.

Some of the confusion disappears if one takes another approach to Rosenthal and Jacobson's statistical analysis. In the Oak School experiment each child in each of 18 classrooms was used either as an "experimental" or a "control" subject in the study. Clearly, this was the easiest and most natural way for the researchers to obtain a reasonably large sample of children. But children affect one another, and many things teachers and children do can affect classroom atmosphere; so there is reason to suspect that each child's test gains cannot be seen as an independent observation for statistical purposes. If the assumption of independence were, in fact, not tenable, we could conclude that they overestimated the degree of statistical stability of much of their data. Accordingly, we reanalyzed some of their major results using the analysis of variance to dis-

cover whether a child's performance gain could be partially accounted for by the particular classroom he happened to be in, independent of his grade level or whether he was in the experimental or the control group. As we suspected, the effect of classroom membership is very stable in the analyses we did ($p < .001$), which means that strictly speaking, the proper "unit of analysis" for the study is the classroom average rather than the individual child (see the footnotes on pp. 95, 96, 118, 119, 145, and 146). Looking at the data this way, many of the results that were statistically significant using the analysis of variance as Rosenthal and Jacobson used it do not reach significance when the more usual and, in fact, more proper test is applied.

On the other hand, even if we confine Rosenthal and Jacobson to these more conservative statistical techniques, we find that the most important results of their study remain statistically stable. Further, these results are reasonably easy to interpret consistently. First, it remains clear that a semester after the teachers had been given information about the 'spurters' in their classrooms, the children designated as spurters showed somewhat greater improvement on the intelligence test than did the control children. This overall effect of teachers' expectancies became much greater at the end of one year but a good deal less marked (and no longer statistically stable) at the end of two years. By and large, these effects were more pronounced in the first and second grades than they were in the later grades. Second, the experimental children tended to show greater gains in reading grades on their report cards and were judged by their teacher to be, on the average, more intellectually curious than the control children. Finally, the average degree of improvement shown by the experimental children in a classroom was strongly associated with the average degree of improvement shown by the control children in the same classroom: if the experimental children did especially well, so did the control children. So it would seem

them, and weigh them again, the changes in their weight would turn out to be highly significant statistically, though the thimblefuls of gasoline would still have contributed little, relatively speaking, to the total weight of the automobiles. In assessing the importance of a variable, then, one must take into account not only the stability and the apparent magnitude of its effects, but also the significance of a result of that magnitude in the total context of the research.

These points aside, *Pygmalion in the Classroom* is an important and thought-provoking book; anyone concerned with the problems and practices of education should certainly take the trouble to read it. The study provides a perfectly satisfactory demonstration that the teacher expectation effects hypothesized do indeed take place. Though the Oak School experiment is not as sophisticated as it might be, it was done in a 'natural' setting (which, incidentally, usually makes it difficult to do elegant research) rather than under the more artificial circumstances of the laboratory. The fact that the effect can be demonstrated in the actual classroom under very ordinary conditions is convincing. That the effect appears to be quite strong and stable in spite of the subtlety and simplicity of the experimental induction is especially dramatic. The results of the experiment fairly demand that much serious attention, thought, and research be devoted to the effect on children of the beliefs and attitudes held about them by school administrators, supervisors, and teachers. It also points up the crucial importance of conducting research on just *how* people's expectations of children become realities. It is possible that we will learn to change teachers' behavior toward children before we learn how to change their attitudes toward them, and thus their expectations of them. It is on this last general consideration —how the expectation of the teacher might have led to modifications of her pupils' performance and classroom behavior—that Rosenthal and Jacobson

significance when the more usual and, in fact, more proper test is applied.

On the other hand, even if we confine Rosenthal and Jacobson to these more conservative statistical techniques, we find that the most important results of their study remain statistically stable. Further, these results are reasonably easy to interpret consistently. First, it remains clear that a semester after the teachers had been given information about the 'spurters' in their classrooms, the children designated as spurters showed somewhat greater improvement on the intelligence test than did the control children. This overall effect of teachers' expectancies became much greater at the end of one year but a good deal less marked (and no longer statistically stable) at the end of two years. By and large, these effects were more pronounced in the first and second grades than they were in the later grades. Second, the experimental children tended to show greater gains in reading grades on their report cards and were judged by their teacher to be, on the average, more intellectually curious than the control children. Finally, the average degree of improvement shown by the experimental children in a classroom was strongly associated with the average degree of improvement shown by the control children in the same classroom: if the experimental children did especially well, so did the control children. So it would seem that the general classroom 'atmosphere' generated by all concerned has something to do with what happens. The second and third set of results are of particular interest because they provide us with clues as to how some of the classroom expectancy effects observed in the Oak School study might have taken place, a consideration to which we shall return.

Let us note what we consider to be some of the book's other weaknesses. First, Rosenthal and Jacobson speak repeatedly of "IQ gains," thus implying that changes occurred in children's intellectual "capacity" or intellective functioning as this is usually understood by the layman. Such a claim would, of course, be unjustified. A discussion of the nature of intelligence and theories about its measurement is not possible within the scope of this paper. Suffice it to say that a change in IQ score can be interpreted simply as a change in performance on a test for which age-norms exist. Since rather little normative information is available (especially on younger children) for the intelligence test used in their research, the IQ gains reported by Rosenthal and Jacobson must be interpreted with particular caution.

Second, there is implicit confusion throughout the book about the difference between the "strength" of a variable and the level of statistical significance associated with its effects. Though the two things often go together, it is important to note that a statistical significance level is an index of the probable stability of a result, and not of its magnitude. A very weak effect that is extremely stable will be detected at a high level of statistical significance even though it may not be very meaningful in practical terms. If, for example, we were to weigh a large number of automobiles on a very sensitive scale, add a thimbleful of gasoline to the gas tanks of most of

of education should certainly take the trouble to read it. The study provides a perfectly satisfactory demonstration that the teacher expectation effects hypothesized do indeed take place. Though the Oak School experiment is not as sophisticated as it might be, it was done in a 'natural' setting (which, incidentally, usually makes it difficult to do elegant research) rather than under the more artificial circumstances of the laboratory. The fact that the effect can be demonstrated in the actual classroom under very ordinary conditions is convincing. That the effect appears to be quite strong and stable in spite of the subtlety and simplicity of the experimental induction is especially dramatic. The results of the experiment fairly demand that much serious attention, thought, and research be devoted to the effect on children of the beliefs and attitudes held about them by school administrators, supervisors, and teachers. It also points up the crucial importance of conducting research on just *how* people's expectations of children become realities. It is possible that we will learn to change teachers' behavior toward children before we learn how to change their attitudes toward them, and thus their expectations of them. It is on this last general consideration —how the expectation of the teacher might have led to modifications of her pupils' performance and classroom behavior—that Rosenthal and Jacobson are weakest in their analysis. Though they do speculate about some aspects of the problem, the heart of the matter remains untouched. In short, they have shown us that teachers' expectations of their students' performances have definite consequences for these (subsequent) performances. But they have not shown us how this process works.

In the remainder of our article, we propose to do some more or less systematic speculating about how the Oak School teachers may have fulfilled the researchers' prophecies. We shall begin by arguing that expectation leads to selectivity of attention, perception, and response, and end by discussing just how increases in interpersonal "warmth" and encouragement might actually lead to superior learning and performance.

Some Psychological Effects of Expectation

The study of the influence of expectation upon thinking and behavior has been of interest to psychologists for many years in a variety of contexts. There are literally hundreds of studies that are relevant to the notion that a person's attitude, set, or expectation will affect his perceptions and responses. One tentative conclusion that can be drawn from these studies—of particular interest in thinking about expectancy in the classroom—is the following: a person is more likely to perceive a barely perceptible stimulus if he expects it than if he does not. For example, if an experimental subject is given a list from which words are to be presented tachistoscopically, the recognition threshold for words on the list (i.e., the words he is expecting) will be lower than for those not on the list (Neisser, 1954). A related study is reported by Bruner and Minturn (1955). When shown an ambiguous stimulus which resembled both a letter and a number, for example *B* and

13, subjects who were told to expect letters saw *B* and those who were told to expect numbers reported seeing *13*. Neither of these results is surprising, possibly because they both confirm what we often observe in our everyday activities—that we are likely to see what we expect to see, and that we tend to interpret ambiguous events in such a way as to confirm our own predictions. Similarly, if a teacher expects to see something, she is likely to find evidence of its occurrence sooner or later.

Further, that a person's expectations exert a powerful influence on the behavior of the people with whom he interacts is a well-documented phenomenon. (A review of literature on this point is contained in Rosenthal, 1966.) In the laboratory, for example, if a group of experimenters is told to expect that their subjects will most likely perform well on a particular task, and another group is told to expect their subjects to perform poorly, the subjects in the former group will tend to have significantly higher scores than those in the latter. Expectation effects have even been demonstrated when the experimental subjects were animals (Rosenthal, 1966, also reported in *Pygmalion in the Classroom).* Strupp and Luborsky (1962) make the point that a therapist's expectations about the prognosis of treatment may have much to do with its actual outcome. The teacher, similarly, may have a personal stake in seeing what she expects to see in the classroom.

If you treat an individual as he is, he will stay as he is, but if you treat him as if he were what he ought to be and could be, he will become what he ought to

entirely in advance by previously existing attitudes and beliefs. Expectation may change, vary, increase, or decrease, depending on the nature of an ongoing social interaction and the reciprocal influence that two people have upon one another. In the classroom, for example, it is likely that a teacher's behavior toward a particular child will not depend consistently on previously existing beliefs about what he can or will do but may be modified by what occurs between them.

In order to imagine how such variables might affect an interaction, let us construct an oversimplified situation. Imagine that one of Rosenthal and Jacobson's teachers is told that Johnny Second-Grader, an unremarkable-appearing boy, is about to experience a period of intellectual blossoming. Given a situation in which the teacher is trying to communicate a difficult bit of information to Johnny, she is now more likely to expect an indication of comprehension than if she had been told nothing about potential academic progress. (In fact, she might not otherwise have attempted the communication in the first place.) A minor change in Johnny's behavior at this point, say a nod or smile, may be interpreted as a glimmer of understanding so that, encouraged, the teacher intensifies her efforts to reach him. Consequent subtle changes in the teacher's behavior and attitude, such as alterations in her body posture, tone of voice, perceived interest, facial expression, or verbal praise, may similarly interest and encourage Johnny, leading to his increased motivation and attention, and finally to the reward of mastering something new—as well as the fulfillment of his teacher's expectation. The point is that psychological expectation may have a catalytic effect, evoking interactions and leading to events that depend upon much more than the effects of expectation alone, but which might not, and indeed probably would not, have occurred without the belief that such events were possible.

behavior of the people with whom he interacts is a well-documented phenomenon. (A review of literature on this point is contained in Rosenthal, 1966.) In the laboratory, for example, if a group of experimenters is told to expect that their subjects will most likely perform well on a particular task, and another group is told to expect their subjects to perform poorly, the subjects in the former group will tend to have significantly higher scores than those in the latter. Expectation effects have even been demonstrated when the experimental subjects were animals (Rosenthal, 1966, also reported in *Pygmalion in the Classroom).* Strupp and Luborsky (1962) make the point that a therapist's expectations about the prognosis of treatment may have much to do with its actual outcome. The teacher, similarly, may have a personal stake in seeing what she expects to see in the classroom.

> *If you treat an individual as he is, he will stay as he is, but if you treat him as if he were what he ought to be and could be, he will become what he ought to be and could be.*
>
> *Goethe*

It appears that a person who is in a position to exercise over another the subtle interpersonal influence we have been talking about may do so without being aware either of the content of the message he transmits or the ways he transmits it. The recipient of influence may be equally unaware that any transaction other than the obvious overt one is taking place; he may well do what is expected of him without even realizing that a demand on him is being made. It also appears likely that some people are more effective 'covert persuaders' than others. Such things as the physical appearance, confidence, warmth, friendliness, amount of experience or competence, interest, and status of the persuader seem to affect the extent to which expectancies are influential (see Rosenthal, 1966). Some of these personal characteristics are also related to a person's effectiveness as an overt persuader. So it seems plausible that some teachers may be better covert persuaders in the classroom than others. It is also possible that the teacher who is warm, friendly, and sure of herself is a better covert persuader than her less friendly and less competent colleagues, and that the highly skilled and successful teacher is also good at helping make her prophecies come true.

The manner in which one person influences another is not determined

us construct an oversimplified situation. Imagine that one of Rosenthal and Jacobson's teachers is told that Johnny Second-Grader, an unremarkable-appearing boy, is about to experience a period of intellectual blossoming. Given a situation in which the teacher is trying to communicate a difficult bit of information to Johnny, she is now more likely to expect an indication of comprehension than if she had been told nothing about potential academic progress. (In fact, she might not otherwise have attempted the communication in the first place.) A minor change in Johnny's behavior at this point, say a nod or smile, may be interpreted as a glimmer of understanding so that, encouraged, the teacher intensifies her efforts to reach him. Consequent subtle changes in the teacher's behavior and attitude, such as alterations in her body posture, tone of voice, perceived interest, facial expression, or verbal praise, may similarly interest and encourage Johnny, leading to his increased motivation and attention, and finally to the reward of mastering something new—as well as the fulfillment of his teacher's expectation. The point is that psychological expectation may have a catalytic effect, evoking interactions and leading to events that depend upon much more than the effects of expectation alone, but which might not, and indeed probably would not, have occurred without the belief that such events were possible.

We have discussed something about the effects that a person's expectations may have on his own behavior and thinking and on the behavior and thinking of others. It seems very clear also that the nature of the social structure itself often has similar effects. One of the findings of a study of the British school system reported by Hilde Himmelweit and Judy Wright (1967) involved a comparison of two schools with different policies regarding assignment to tracks (or streams, as they are referred to in England). In one school, assignment to a stream was based on ability. In the other, assignment was based on criteria other than intellectual ability. Yet the effect of stream placement on further academic advancement and final performance was identical for the two schools. Thus, the effect of streaming seemed to be a more powerful determinant of performance than were the attributes that led to the pupils' initial placement in streams; streaming turns out to be a potent 'leveler' in Britain. This result can be interpreted as indicating that the meaning for an individual of being allocated to a particular stream, and the influence exerted by the experience of stream membership, can be major determinants of his progress. The expectation for a person's behavior that is implicit by virtue of his place in the social structure may exert a powerful influence over what he does, in that he will not only respond to the expectation but may also help to create an environment within which it will be fulfilled.

Let us be very clear. As common and powerful as these so-called 'expectation' effects seem to be, virtually nothing is known about how they are communicated. We know that the classroom teacher expects Johnny to be a 'spurter' this year, and we know that she therefore probably wants him to improve. We can guess that she does something which communi-

cates itself to Johnny, and it seems clear that Johnny responds to whatever it is she does. But we don't know what she does, or exactly what Johnny picks up. Furthermore, how does Johnny's 'receiving the message,' whatever it might be, actually lead to his improvement? These questions are just beginning to come under the scrutiny of research psychologists. As a basis for some further speculation, we shall discuss an ingenious experiment conducted in the Social Psychology Laboratory at Teachers College by Donnah Canavan.

An Experimental Study of the Development of "Reward-Orientation"

Canavan distinguishes between two general styles that can be used by a person to attempt to maintain satisfaction in his environment: reward orientation and cost orientation (see Thibaut and Kelley, 1959). A reward-oriented individual is described as one who attempts to act on his environment to obtain satisfying things for himself. In viewing what the world may have in the offing for him, he pays more attention to potential rewards than he does to potential costs. He is usually confident, outgoing, and relatively skillful at controlling the environment. A cost-oriented individual, on the other hand, maintains his satisfaction by keeping his standards for satisfaction relatively low and by using his energy to avoid costs and failures; he attends more to the potential costs in the environment than to the rewards that may lie in it. He tends to feel unconfident, constricted, and relatively powerless to influence the environment. Often he has fewer skills for handling the environment than does the reward-oriented person.

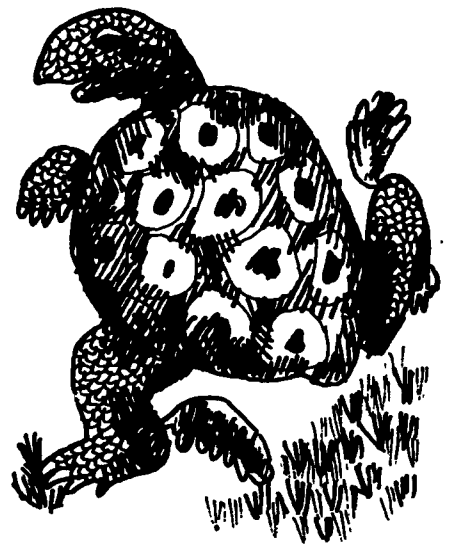
Canavan speculates that the development of reward or cost orientation may be a kind of self-fulfilling prophecy that begins with the social labeling of successes and failures by 'powerful others,' such as parents and teachers, who are usually in control of information about criteria for evaluating behavior and outcomes. She suggests, in other words, that people are taught by others how to look at and evaluate their experiences.

Let us suppose that an individual is given warm social praise by a powerful other for his successes, and no particular attention is paid to his unsuccessful ventures. For this person, the rewarding aspects of the world are likely to became particularly salient, as compared to the costly aspects of

ment to obtain satisfying things for himself. In viewing what the world may have in the offing for him, he pays more attention to potential rewards than he does to potential costs. He is usually confident, outgoing, and relatively skillful at controlling the environment. A cost-oriented individual, on the other hand, maintains his satisfaction by keeping his standards for satisfaction relatively low and by using his energy to avoid costs and failures; he attends more to the potential costs in the environment than to the rewards that may lie in it. He tends to feel unconfident, constricted, and relatively powerless to influence the environment. Often he has fewer skills for handling the environment than does the reward-oriented person.

Canavan speculates that the development of reward or cost orientation may be a kind of self-fulfilling prophecy that begins with the social labeling of successes and failures by 'powerful others,' such as parents and teachers, who are usually in control of information about criteria for evaluating behavior and outcomes. She suggests, in other words, that people are taught by others how to look at and evaluate their experiences.

Let us suppose that an individual is given warm social praise by a powerful other for his successes, and no particular attention is paid to his unsuccessful ventures. For this person, the rewarding aspects of the world are likely to became particularly salient, as compared to the costly aspects of the world. He may come to regard the costs he experiences as simply instances of nonreward. If, on the other hand, an individual is warned and criticized about failures and relatively little attention is paid to his successes, then the costs he experiences would become particularly salient for him. Since his primary concern would be to avoid the social costs associated with his failures, he might come to regard rewarding circumstances simply as instances of successful failure-avoidance. The reward-oriented individual would be expected to develop higher standards for performance than would the cost-oriented person, who would tend to confine his aspirations to relatively low levels of outcomes in order to maintain a high probability of avoiding failure. Since an individual's orientation to rewards or to costs affects both his decisions and his actions, the consequences of maintaining a particular orientation are likely to be important and far-reaching. Canavan's experiment attempted to initiate a 'reward-orientation' cycle in one group of children and a 'cost-orientation' cycle in another group, and to observe the development and the effects of the cycles that were initiated.



The experimental subjects were 64 fifth and sixth grade girls in two New York suburban schools. The girls performed a series of 'decoding' exercises which had eight levels of difficulty. In the exercises, subjects translated symbols into numbers using a code key, and then summed the numbers to arrive at an answer, which was then further summed to arrive at a final one or two-digit number. The level of difficulty of an exercise was varied by changing the proportion of numbers to be decoded.

The subject did two sets of exercises. In the first set, which was presented to her as "practice," she did 20 exercises, ten at the easiest level (two symbols) and ten at the most difficult (16 symbols). At the end of each of 16 of

the exercises, she was told what the correct answer was; on four, she received no information about whether her answer was correct. Though the subject did not realize it, the experimenter told her what the 'correct' answer was according to a pre-programmed schedule, so that every subject in each of the experimental treatments believed she got six easy problems right and two wrong, and two difficult problems right and six wrong. During the practice exercises, training which attempted to begin the process of reward or cost orientation was given as follows; there were four "treatments:"

> 1. In the Reward-Success Only treatment, the experimenter paid particular attention to the subject's "correct" answers, saying such things to her as "That's right! That's very good!" in a warm and encouraging fashion when her answer to a problem, regardless of difficulty, was "c rect." No comments were made when the subject's answers were "incorrect."
>
> 2. In the Punish-Failure Only treatment, no attention was paid to the subject's "correct" answers, and her "incorrect" answers were emphasized by the experimenter's remarks. On the "incorrect" exercises, the experimenter said such things as, "That one's wrong—that's not too good."
>
> 3. In the Reward-Success and Punish-Failure treatment, both "correct" and "incorrect" exercises were responded to by the experimenter, exactly as described above.
>
> 4. In the Neither Reward-Success nor Punish-Failure condition, the experimenter responded only with the "correct" answer to the exercises, and said nothing else to the subject.

After the practice exercises, the subject responded to a questionnaire having to do with how she felt about various aspects of the task, and then turned her attention to the next set of exercises. Here she was allowed to choose the level of difficulty she wanted before she did each of ten exercises. She was told she would get "points" for each of the exercises she got correct; the number of points varied from two to 16, depending on the level of difficulty of the problem she chose.

After each of the ten exercises which the subject now completed, the experimenter told her only what the actual correct answer was, adding no [illegible] to observe in these ten exercises some

crepancy between how well subjects expected to perform and how well they felt they had to perform in order to be minimally satisfied was greatest for subjects in the Reward-Success Only treatment, and least for subjects in the Punish-Failure Only treatment. It seems reasonable to speculate that feeling some latitude to experiment, to risk new ways of doing things or new levels of task difficulty, may be useful for the development of new skills as well as feelings of mastery and self-confidence. Indeed, subjects in the Reward-Success Only treatment tended to be more confident that they would do well and less bothered by their failures than subjects in the other treatments.

Let us return now to the classroom and to the teacher who expects to see startling improvement in the performance of a particular pupil. As we have argued above, if she expects a pupil to begin to improve, she may be avidly watching for signs of improvement while ignoring the pupil's usual inadequacies or failures. If she should see something that indicates improvement, she might be especially quick to reward it by her special attention and by her excitement at seeing her expectations begin to be fulfilled. This in turn might give the child new interest in this kind of performance, and might spur him on to new attempts to fulfill an expectation that he might now begin to perceive. Since the teacher is no longer paying as much attention to the child's failures, the child may now feel new room to grow, and indeed might grow with his burgeoning confidence about new learning and new power over his environment. As the child's performance improves, his teacher's standards for his performance may become higher; thus, what Rosenthal and Jacobson term the "benign circle" might develop. And here is an added dividend: a teacher who spends relatively more time rewarding success in the classroom, and therefore relatively less time punishing failure, may be improving the learning environment not only for the few children of whom she expects new things, but for all the children in the room—as is suggested by the results in Rosenthal and Jacobson described above.

Our discussion in this section has been very speculative; certainly, many links are missing in the chain that connects what the teacher expects the student to be able to do and what the student becomes able to do. And,

were emphasized by the experimenter's remarks. On the "incorrect" exercises, the experimenter said such things as, "That one's wrong—that's not too good."

3. In the Reward-Success and Punish-Failure treatment, both "correct" and "incorrect" exercises were responded to by the experimenter, exactly as described above.

4. In the Neither Reward-Success nor Punish-Failure condition, the experimenter responded only with the "correct" answer to the exercises, and said nothing else to the subject.

After the practice exercises, the subject responded to a questionnaire having to do with how she felt about various aspects of the task, and then turned her attention to the next set of exercises. Here she was allowed to choose the level of difficulty she wanted before she did each of ten exercises. She was told she would get "points" for each of the exercises she got correct; the number of points varied from two to 16, depending on the level of difficulty of the problem she chose.

After each of the ten exercises which the subject now completed, the experimenter told her only what the actual correct answer was, adding no other comments. Thus, we are able to observe in these ten exercises some of the effects of the earlier training method that the experimenter used.

Canavan was interested in the level of risk or difficulty that subjects in the various treatments would choose, the number of actual errors they made, the number of points they earned, and how they felt about the task.*

As Canavan predicted, subjects in the Reward-Success Only treatment chose more difficult problems, made fewer errors (in spite of having done more difficult problems), and earned more points than did subjects in any of the other conditions. The Punish-Failure Only subjects fell at the other extreme on all the measures, and subjects in the other two treatments fell between the two extreme treatments on the measures. Further, the dis-

* *Table 1. Summary of Some Results of the Canavan Experiment*[1]

MEASURES	EXPERIMENTAL TREATMENT (TYPE OF SOCIAL TRAINING)			
	Reward Success Only	*Reward Success and Punish Failure*	*Neither Reward Success nor Punish Failure*	*Punish Failure Only*
Average difficulty level chosen in 10 free choice exercises	5.8	5.2	5.2	4.3
Average number of errors made in 10 free choice exercises	2.2	2.8	2.4	3.4
Average number of points obtained in 10 free choice exercises	92.1	76.4	77.0	56.1
Average Index of Perceived "Room to Maneuver"[2]	2.2	1.2	1.8	0.6

[1] *The four sets of means presented have been subjected to preliminary statistical analyses. In each case the finding that the Reward Success Only and Punish Failure Only means fall at the extremes while the other two means fall between the extremes is supported at conventionally acceptable significance levels, either by tests of ordered means (Page's L test) or analysis of variance, or both.*

[2] *For this index, each subject was asked (a) the number of exercises (out of ten) on which she expected to be correct, and (b) the smallest number she could get correct and still remain satisfied with her performance; (b) was then subtracted from (a).*

tion and by her excitement at seeing her expectations begin to be fulfilled. This in turn might give the child new interest in this kind of performance, and might spur him on to new attempts to fulfill an expectation that he might now begin to perceive. Since the teacher is no longer paying as much attention to the child's failures, the child may now feel new room to grow, and indeed might grow with his burgeoning confidence about new learning and new power over his environment. As the child's performance improves, his teacher's standards for his performance may become higher; thus, what Rosenthal and Jacobson term the "benign circle" might develop. And here is an added dividend: a teacher who spends relatively more time rewarding success in the classroom, and therefore relatively less time punishing failure, may be improving the learning environment not only for the few children of whom she expects new things, but for all the children in the room—as is suggested by the results in Rosenthal and Jacobson described above.

Our discussion in this section has been very speculative; certainly, many links are missing in the chain that connects what the teacher expects the student to be able to do and what the student becomes able to do. And, surely, there is not just one thing happening in the entire process, but many things. The phenomenon of subtle interpersonal influence guiding progress in the classroom is as complex as it is fascinating. Our plea is merely that these matters, adumbrated by *Pygmalion in the Classroom*, are researchable now, at our present level of research sophistication, and are especially deserving of a major laboratory and field research effort at this time.

References

1. *J. S. Bruner, and A. L. Minturn. "Perceptual Identification and Perceptual Organization."* Journal of Genetic Psychology, *1955, 53, pp. 21–28.*
2. *Donnah Canavan. Development of Individual Differences in the Perception of Value and Risking-taking Styles. Ph.D. dissertation. New York: Teachers College, Columbia University (in preparation).*
3. *Hilde Himmelweit, and Judy Wright. "The School System, Social Class, and Attainment After School." Paper presented at the Annual Conference of the British Psychological Society, 1967.*
4. *U. Neisser. "On Experimental Distinction Between Perceptual Process and Verbal Response."* Journal of Experimental Psychology, *1954, 47, pp. 399–402.*
5. *R. Rosenthal.* Experimenter Effects in Behavioral Research. *New York: Appleton-Century-Crofts, 1966.*
6. *H. H. Strupp, and L. Luborsky, Eds.* Research in Psychotherapy. *Washington, D.C.: American Psychological Association, 1962.*
7. *J. W. Thibaut, and H. H. Kelley.* The Social Psychology of Groups. *New York: Wiley and Sons, 1959.*

Peter Gumpert is an assistant professor in the doctoral program in social psychology at Teachers College, Columbia University. His wife, Carol Gumpert, is a clinical psychologist in the department of psychiatry of the Albert Einstein College of Medicine.