Pygmalion in the Classroom*

by Robert Rosenthal & Lenore Jacobson

In 1965 the authors conducted an experiment in a public elementary school, telling teachers that certain children could be expected to be "growth spurters," based on the students' results on the Harvard Test of Inflected Acquisition. In point of fact, the test was nonexistent and those children designated as "spurters" were chosen at random. What Rosenthal and Jacobson hoped to determine by this experiment was the degree (if any) to which changes in teacher expectation produce changes in student achievement.

Robert Rosenthal is professor of social psychology at Harvard University. Lenore Jacobson is an elementary school principal in the South San Francisco Unified School District.

The basic question to be answered in this chapter is whether in a period of one year or less the children of whom greater intellectual growth is expected will show greater intellectual growth than the undesignated control-group children. There are also four important subsidiary questions. If there were some advantages to a child whose teacher had favorable expectations for his intellectual development, would these expectancy advantages be greater for:

1. Children in the lower grades or higher grades?
2. Children in the fast track, or medium track, or slow track?
3. Children of one sex rather than the other?
4. Children of minority group or nonminority group status?

THE MAJOR VARIABLES

Age

The folk knowledge of our culture, current theories of human development, especially psychoanalytic theory, and the work of the developmental and experimental psychologists and of the ethologists are in agreement on the importance of age as a factor in determining the degree to which an organism can be shaped, molded, or influenced (Scott, 1962). In general, the younger the organism, the greater is thought to be the degree of susceptibility to social influence. In his classic monograph, Coffin (1941) concluded that influenceability increased from infancy to ages seven to nine but decreased after that. More recently in a summary of the evidence bearing on overt social influence on children, Stevenson (1965) reported the greater influenceability of five-year-olds than twelve-year-olds, a finding consistent with Coffin's summary. Both Coffin and Stevenson were writing about more overt social influence than the subtle, unintended influence of teachers' prophecies. Still, it would be interesting to know whether influence processes of a more subtle, unintended form would also show younger children to be the more susceptible.

Ability

We are also interested in learning whether the children of the three tracks differ in the degree to which they profit from the teachers' favorable expectations. In the case of ability, however, the literature is not so helpful in telling us what we might find. Stevenson (1965) suggested that susceptibility to social influence may not be too contingent on the child's intellectual status, and we know that the three tracks differ considerably in average IQ. One of the most recent discussions of intellectual gains is by Thorndike (1966) who reports that there are only modest correlations between initial intellectual status and changes in intellectual status. In the present research, in any case, we are not so much interested in gains *per se* but rather in the excess of gain that might be shown by the "special" children over the "ordinary" undesignated children. In short, we are interested in differences among the tracks in the degree of expectancy advantage that may be found, but we hardly know what to expect. The matter is further complicated by the fact that the other two variables in which we are interested, sex and minority group status, are not independent of track placement. In the last chapter we saw that boys tend to overpopulate the slow track relative to girls who tend to overpopulate the fast track. Mexican children, Oak School's minority group, tend to overpopulate the slow track and underpopulate the fast track.

Sex

Whether boys or girls are the more susceptible to social influence processes depends on whether the influencer is male or female (Stevenson, 1965). Since the overwhelming majority of Oak School's teachers are females, the findings from research with lady influencers interest us most. Those findings summarized by Stevenson (1965), suggest that boys should be the more susceptible to social influence. As in the case of the children's age, however

the social influence processes employed were neither unintended nor very subtle. Effects of teachers' expectations are likely to be both.

Minority Group Status

The reasons for our interest in the variable of minority-group status need little justification. So much of the literature on the disadvantaged child focuses on the minority-group child that "disadvantaged" almost means "minority group." One of the best known publications dealing with the disadvantaged is called *Youth in the Ghetto*. We shall be especially interested, then, if expectancy advantages occur at all, in whether they benefit minority-group children more or less than nonminority-group children.

At Oak School the minority-group child is Mexican. The definition of a minority-group child in this research, however, was more stringent than simply whether the name was Mexican. To qualify as a "minority-group child," either the child himself or his parents had to come from Mexico, Spanish had to be spoken at home, and the child had to be present for the administration of certain procedures. These procedures, in connection with another study (Jacobson, 1966), included administration of an IQ test in Spanish, a test of reading ability, and the taking of photographs of the child himself. Within this sample of Mexican minority-group children there were variations in how "Mexican" each child looked. A group of ten teachers with no connection to Oak School or its children rated each photograph on "how Mexican the child looked." The definition of how clearly Mexican a child "really" looked was the average rating of all ten teachers. These ratings were highly reliable. The average rating of the same children by the teachers of Oak School was correlated .97 with the ratings of the judges who were not associated with Oak School.

INTELLECTUAL GROWTH

Expectancy Advantage by Grades

The bottom row of Table 1 gives the over-all results for Oak School. In the year of the experiment, the undesignated control-group children gained over eight IQ points while the experimental-group children, the special children, gained over twelve. The difference in gains could be ascribed to chance about 2 in 100 times ($F = 6.35$).

The rest of Table 1 and Figure 1 show the gains by children of the two groups separately for each grade. We find increasing expectancy advantage as we go from the sixth to the first grade; the correlation between grade level and magnitude of expectancy advantage ($r = -.86$) was significant at the .03 level. The interaction effect, or likelihood that at different grades there were significantly greater expectancy advantages, was significant at the .07 level ($F = 2.13$). (Interactions, however, are not sensitive to the ordering of differences unless one makes them so with further statistical efforts; that is, the p of .07 is conservative.)

In the first and second grades the effects of teachers' prophecies were dramatic. Table 1 shows that, and so does Table 2 and Figure 2. There we find the percentage of experimental- and control-group children of the first two grades who achieved various amounts of gain. In these grades about every fifth control-group child gained twenty IQ points or more, but of the special children, nearly every second child gained that much.

So far we have told only of the effects of favorable expectancies on total IQ, but Flanagan's TOGA yields separate IQs for the verbal and reasoning spheres of intellectual functioning. These are sufficiently different from each other so it will not be redundant to give the results of each. In the case of verbal IQ the control-group children of the entire school gained just less than eight points, and the special children gained just less than ten, a difference that could easily have arisen by chance. The interaction term was not very significant ($p < .15$) so that we can not conclude greater expectancy advantage at some grade levels than at others. But we do have a special interest now in the first and second graders, and it will do no harm

Table 1
MEAN GAIN IN TOTAL IQ AFTER ONE YEAR BY EXPERIMENTAL- AND CONTROL-GROUP CHILDREN IN EACH OF SIX GRADES

	Control		Experimental		Expectancy Advantage	
Grade	N	Gain	N	Gain	IQ Points	One-Tail $p < .05$[a]
1	48	+12.0	7	+27.4	+15.4	.002
2	47	+ 7.0	12	+16.5	+ 9.5	.02
3	40	+ 5.0	14	+ 5.0	− 0.0	
4	49	+ 2.2	12	+ 5.6	+ 3.4	
5	26	+17.5 (−)	9	+17.4 (+)	− 0.0	
6	45	+10.7	11	+10.0	− 0.7	
Total	255	+ 8.42	65	+12.22	+ 3.80	.02

[a] *Mean square within treatments within classrooms = 164.24.*

Figure 1
GAINS IN TOTAL IQ IN SIX GRADES

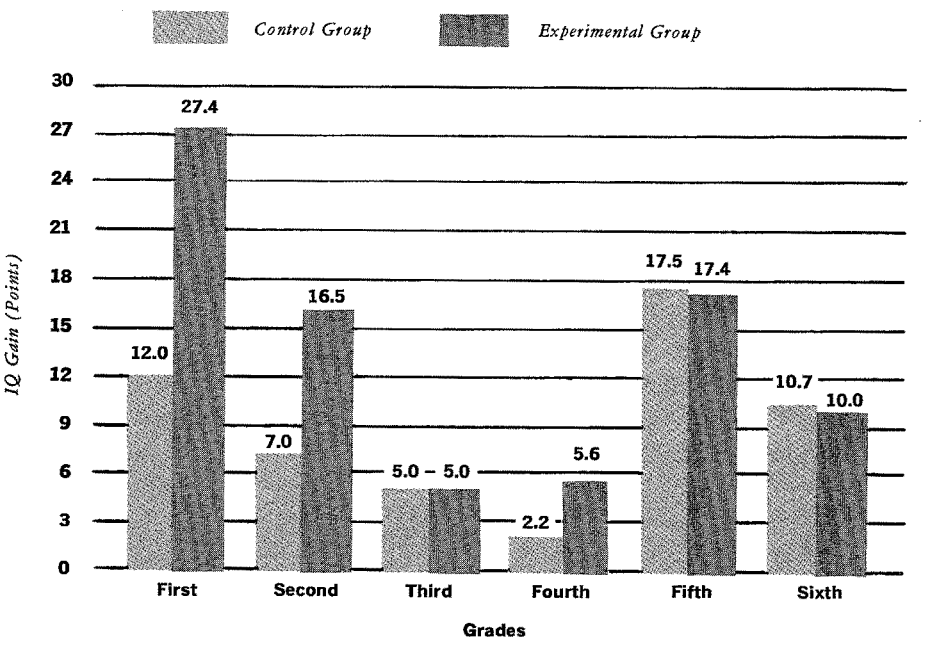

Table **2**
PERCENTAGES OF FIRST AND SECOND GRADERS GAINING AT LEAST TEN, TWENTY, OR THIRTY TOTAL IQ POINTS

IQ Gain at Least	*Control N = 95*	*Experimental N = 19*	*One-Tail p of Difference*
10 points[a]	**49%**	**79%**	**.02**
20 points[b]	**19%**	**47%**	**.01**
30 points	**5%**	**21%**	**.04**

[a] *Includes children gaining twenty and thirty points or more.*
[b] *Includes children gaining thirty points or more.*

Figure **2**
PERCENTAGES OF FIRST AND SECOND GRADERS GAINING TEN, TWENTY, OR THIRTY TOTAL IQ POINTS

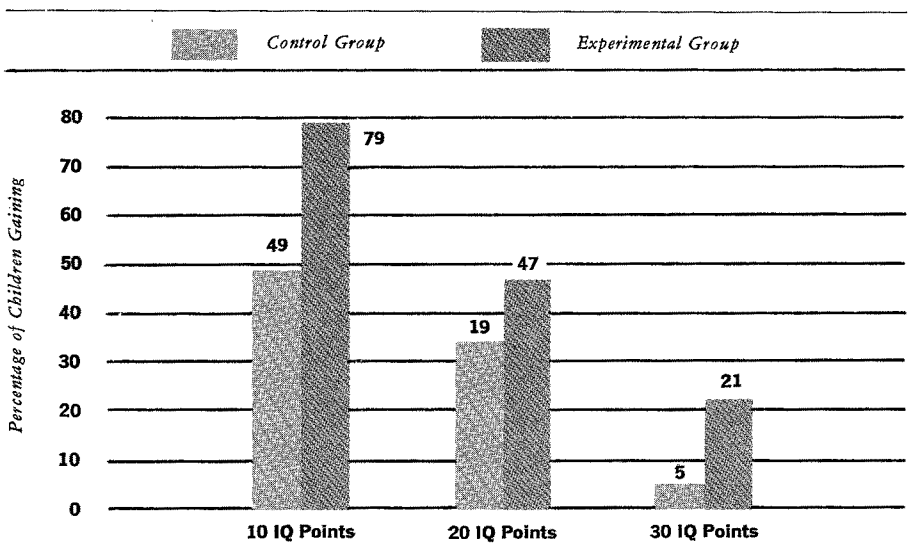

Table **3**
MEAN GAIN IN VERBAL IQ AFTER ONE YEAR BY EXPERIMENTAL- AND CONTROL-GROUP CHILDREN IN GRADES ONE-TWO AND THREE-SIX

	Control		*Experimental*		*Expectancy Advantage*	
Grades	*N*	*Gain*	*N*	*Gain*	*IQ Points*	*One-Tail p < .05*[a]
1–2	**95**	**+4.5**	**19**	**+14.5**	**+10.0**	**.02**
3–6	**174**	**+9.6**	**49**	**+ 8.0**	**− 1.6**	
Total	**269**	**+7.79**	**68**	**+ 9.85**	**+ 2.06**	

[a] *Mean square within = 316.40.*

Table **4**
MEAN GAIN IN REASONING IQ AFTER ONE YEAR BY EXPERIMENTAL- AND CONTROL-GROUP CHILDREN IN GRADES ONE-TWO AND THREE-SIX

	Control		*Experimental*		*Expectancy Advantage*	
Grades	*N*	*Gain*	*N*	*Gain*	*IQ Points*	*One-Tail p < .05*[a]
1–2	**95**	**+27.0 (−)**	**19**	**+39.6 (+)**	**+12.7**	**.03**
3–6	**160**	**+ 9.1 (−)**	**46**	**+15.9 (+)**	**+ 6.9**	**.06**
Total	**255**	**+15.73**	**65**	**+22.86**	**+ 7.13**	**.005**

[a] *Mean square within = 666.58.*

to see what happened there in particular. In those combined grades, the control-group children gained 4.5 verbal IQ points, and the special children gained exactly 10 points more, or 14.5. If we may have the t test (even though the interaction effect was not significant) we would find $t = 2.24$, $p < .02$, one-tail.

For grades three through six the control gained 1.6 points more than the experimental group, a difference not nearly significant. Table 3 summarizes these results.

The advantage of favorable expectations showed itself more clearly in reasoning IQ as shown in Table 4. For the school as a whole, the advantage of favorable expectations was a seven point net gain in reasoning IQ ($F^* = 6.98$), and there were no significant differences in the six grades in degree of expectancy advantage. Once again, the younger children benefited most. While we are not especially interested in the magnitude of IQ gain of the control group, it does seem remarkable that the younger children of even the control group should gain so heavily in reasoning IQ. Table 1 shows that control-group children gained substantially in total IQ and not only at the younger ages where we might expect practice effects to be most dramatic. There is no way to be sure about the matter (we shall return to it in a later chapter), but it may be that experiments are good for children even when the children are in the untreated control group.

Expectancy Advantage by Tracks and Sex

None of the statistical tests showed any differences among the three tracks in the extent to which they benefited from teachers' favorable prophecies. That was the case for total IQ, verbal IQ, and reasoning IQ. When the entire school benefited as in total IQ and reasoning IQ, all three tracks benefited; and when the school as a whole did not benefit much, as in verbal IQ, none of the tracks showed much benefit. For all three IQ measures, the tendency was for the middle track, the more average children, to benefit most from being expected to grow intellectually, but the difference could easily have occurred by chance.

In total IQ, girls showed a slightly greater advantage than boys of having been expected to show an intellectual spurt; but to see what really happened we must look at boys' and girls' expectancy advantages for the two subtypes of IQ. Table 5 shows the gains in all three types of IQ by boys and girls of the experimental and control groups. In verbal IQ it was the boys who showed the expectancy advantage (interaction $F = 2.13$, $p = .16$); in reasoning IQ it was the girls who showed the advantage, and it was dramatic in size (interaction $F = 9.27$, $p = .003$). Just why that should be is not at all clear. On the pretest, boys had shown a higher verbal IQ than girls (4.4 points), and girls had shown a higher reasoning IQ than boys (8.5 points). Apparently each group profited more from teachers' prophecies in the area of intellectual functioning in which they were already a little advantaged.*

It was mentioned earlier that expectancy advantage was not dependent on placement in any one of the three tracks. That conclusion is modified when we examine expectancy advantages in the three tracks separately for boys and girls. Only for reasoning IQ is there a statistically significant effect (triple interaction $F = 3.47$, $p < .04$). Table 6 shows the excess of gain in reasoning IQ by the experimental over the control boys and girls in each of the three tracks. We already knew that girls showed the greater expectancy advantage in reasoning IQ, and from Table 6 we see that this was significantly more true in the medium track, the track with the more average children.

We knew also that girls are over-represented in the fast track. There are the brighter girls from whom a lot is already expected. The slow track girls tend to be relatively very slow at Oak School, and we know that girls only rarely are placed there, and that they represent a real challenge to Oak School's teachers. Of the middle-track girls there is little to say—

teachers tend to find them uninteresting; pre-existing expectations about their intellectual ability are neither favorable as in the fast track nor very unfavorable and challenging as in the slow track. Perhaps when teachers are given favorable expectations about these children a greater increment of interest results than when expectations are given of girls in the outer tracks. That is a possible explanation of the greater effect in the average track of teachers' favorable expectations for girls' intellectual growth. Why the growth should be in reasoning IQ in particular is not at all clear, but we do know that for the girls in this experiment when there are advantages of teacher prophecies they tend to occur in the reasoning sphere of intellectual functioning.

A pupil's sex turned out to be a factor complicating the amount of expectancy advantage found in the three tracks. Sex also complicated the magnitude of expectancy advantage found in the younger children of the first two grades compared to the older children of the upper four grades. Table 7 shows the number of IQ points by which the gains of the experimental-group children exceeded the gains of the control-group children. These expectancy advantage scores are shown separately for each of the three IQ measures for boys and girls in the lower and upper grades. For total IQ, although the "special" boys of the lower grades did profit from being expected to grow intellectually, the girls of the lower grades gained nearly three times as many IQ points as a function of favorable expectations (triple interaction $F = 2.96$, $p = .09$). For verbal IQ there was no difference between boys and girls at either grade level in the amount of profit from favorable expectations (triple interaction $F < 1$), although, as we learned earlier, boys and girls of the lower grades were helped more than children of upper grades three through six. For reasoning IQ, boys and girls at different grade levels did show very different magnitudes of expectancy advantage. Boys in higher grades performed better in contrast to girls in lower grades who performed better when they were expected to do better (triple interaction $F = 8.14$, $p < .005$). Most of that effect was due to the extraordinary performance of the first- and second-grade girls of the experimental group who gained over forty IQ points more than did the control-group girls of the first and second grade.

To summarize our somewhat complex findings involving pupil's sex as a factor, we may say most simply that girls bloomed more in the reasoning

** This footnote will serve to illustrate the complexity of nature and the need for nonchalancy in the behavioral researcher. Preliminary results of a study conducted with Judy Evans give just the opposite results and with an equally significant probability level. The same basic experiment conducted at Oak School was repeated in two elementary schools located in a small Midwestern town. Unlike Oak School, which drew its pupils from a lower-class community, these schools drew their pupils from a substantial middle-class community. Oak School's student body included a large proportion of minority-group members; the two Midwestern schools did not. The mean pretest total IQ at Oak School was 98, compared to the pretest total IQ of 105 found in these Midwestern schools. Eight months after the teachers had been given the names of their "special" children, retests were administered. The results of the studies at the two schools were sufficiently similar that the results could reasonably be combined. No expectancy advantage was found for either boys or girls as measured by total IQ or verbal IQ. For reasoning IQ, however, the results were opposite to those found at Oak School. Now it was the boys who showed the benefits of favorable teacher expectations. Those who had been expected to bloom gained over sixteen IQ points compared to the less than nine gained by control-group boys. Among the girls it was the control-group children who gained about fifteen IQ points while those of the experimental group gained just over five IQ points. (The interaction F was 9.10, $p < .003$.) In these schools, just as in Oak School, boys had shown higher pretest verbal IQs than girls while girls had shown higher pretest reasoning IQs than boys. Therefore, in these middle-class schools it was not true that each sex benefited most from favorable teacher expectations in those areas in which they were already somewhat advantaged. At the time of this writing there appears to be no ready explanation for this dramatic and very highly statistically significant reversal ($p = .00004$) in the two studies. But now we know for sure that Oak School's results, like the results of all behavioral experiments, are not universal.*

Table **5**
MEAN GAIN IN THREE IQ SCORES AFTER ONE YEAR BY EXPERIMENTAL AND CONTROL BOYS AND GIRLS

	Control		Experimental		Expectancy Advantage	
	N	*Gain*	*N*	*Gain*	*IQ Points*	*One-Tail p < .06*
Total IQ						
Boys	**127**	**+ 9.6**	**32**	**+12.5**	**+ 2.9**	
Girls	**128**	**+ 7.3**	**33**	**+12.0**	**+ 4.7**	**.04**
Verbal IQ						
Boys	**136**	**+ 8.4 (−)**	**34**	**+13.9 (+)**	**+ 5.6**	**.06**
Girls	**133**	**+ 7.2**	**34**	**+ 5.8**	**− 1.4**	
Reasoning IQ						
Boys	**127**	**+19.2**	**32**	**+15.3**	**− 3.9**	
Girls	**128**	**+12.3**	**33**	**+30.2**	**+17.9**	**.0002**

Table **6**
EXCESS OF GAIN IN REASONING IQ BY EXPERIMENTAL OVER CONTROL BOYS AND GIRLS IN THREE TRACKS AFTER ONE YEAR

Track	*Boys*	*Girls*
Fast	**− 2.6**	**+ 9.1**
Medium	**−12.0**	**+42.0[a]**
Slow	**− 0.3**	**+12.5**
Total	**− 3.9**	**+17.9**

[a] *p = .00003, one-tail.*

Table **7**
EXCESS OF GAIN IN THREE IQ SCORES BY EXPERIMENTAL OVER CONTROL BOYS AND GIRLS IN TWO GRADE LEVELS AFTER ONE YEAR

	Boys	*Girls*
Total IQ		
Grades 1–2	**+ 6.1**	**+17.1[b]**
Grades 3–6	**+ 2.3**	**− 0.1**
Verbal IQ		
Grades 1–2	**+10.8[a]**	**+ 9.5**
Grades 3–6	**+ 2.8**	**− 5.8**
Reasoning IQ		
Grades 1–2	**−10.7**	**+40.2[c]**
Grades 3–6	**+ 3.6**	**+10.0[a]**

[a] *p < .05, one-tail (or .10 two-tail).*
[b] *p < .0002, one-tail.*
[c] *p < .00002, one-tail.*

Table **8**
CORRELATIONS BETWEEN MEXICAN FACIAL CHARACTERISTICS AND ADVANTAGES OF FAVORABLE EXPECTATIONS AFTER ONE YEAR

	Boys		*Girls*		*Total*	
	N	*r*	*N*	*r*	*N*	*r*
Total IQ	**7**	**+.70[a]**	**9**	**−.14**	**16**	**+.27**
Verbal IQ	**7**	**+.54**	**10**	**−.11**	**17**	**+.21**
Reasoning IQ	**7**	**+.75[b]**	**9**	**−.01**	**16**	**+.14**

[a] *p = .08, two-tail.*
[b] *p = .05, two-tail.*

sphere of intellectual functioning, and boys bloomed more in the verbal sphere of intellectual functioning when some kind of unspecified blooming was expected of them. Furthermore, these gains were more likely to occur to a dramatic degree in the lower grades. That susceptibility to the unintended influence of the prophesying teacher should be greater in the lower grades comes as no special surprise. All lines of evidence tend to suggest that it is younger children who are the more susceptible to various forms of influence processes. The influence of a teacher holding favorable expectations may not be so very different. Why the boys gained more in verbal IQ when expected to gain intellectually, and why the girls gained more in reasoning IQ is not so easily explained. Earlier we did mention the possibility that children profit more from vague teacher expectations in those spheres of intellectual functioning in which they tend to be slightly advantaged to begin with. In Oak School, the pretest verbal IQs were higher for boys than for girls by over four points; the pretest reasoning IQs were higher for girls than for boys by over eight points.

Expectancy Advantage by Minority-Group Status

In total IQ, verbal IQ, and especially reasoning IQ, children of the minority group were more advantaged by favorable expectations than were the other children though the differences were not statistically significant.

For each of the Mexican children, the magnitude of expectancy advantage was computed by subtracting from his or her IQ gain the IQ gain made by the children of the control group in his or her classroom. The resulting magnitudes of expectancy advantage were then correlated with the "Mexican-ness" of the children's faces. Tables 7-8 show the correlations obtained among Mexican boys and girls when expectancy advantage was defined by total, verbal, and reasoning IQs. For total IQ and reasoning IQ, those Mexican boys who looked more Mexican benefited more from teachers' favorable expectations than did the Mexican boys who looked less Mexican. There is no clear explanation for these findings, but we can speculate that the teachers' pre-experimental expectancies of the more Mexican-looking boys' intellectual performance was probably lowest of all. These children may have had the most to gain by the introduction of a more favorable expectation into the minds of their teachers.

SOME DISCUSSION

The results of the experiment we have described in some detail provide further evidence that one person's expectations of another's behavior may come to serve as a self-fulfilling prophecy. When teachers expected that certain children would show greater intellectual development, those children did show greater intellectual development. For the basic year of the experiment, the self-fulfilling prophecy was in evidence primarily at the lower grade levels; it is difficult to be certain why that was the case. A number of interpretations suggest themselves, and these are not mutually exclusive.

First, younger children are generally regarded as more malleable, less fixed, more capable of change, more subject to the effects of critical periods (Scott, 1962). It may be, then, that the experimental conditions of our experiment were more effective with younger children simply because younger children are easier to change than older ones. (It should be recalled that when we speak here of change we mean it as change relative to control-group change. Table 1 showed that even fifth graders can change dramatically in IQ, but there the change of the experimental-group children was not greater than the change of the control-group children.)

A second interpretation is that younger children within a given school have less well-established reputations within the school. It then becomes more credible to a teacher to be told that a younger child will show intellectual growth. A teacher may "know" an older child much better by reputation and be less inclined to believe him capable of intellectual growth simply on someone else's say-so.

A third interpretation is a combination, in a sense, of the first two. It suggests that younger children show greater gains associated with teachers' expectancies not because they necessarily *are* more malleable but rather because they are believed by teachers to be more malleable.

A fourth interpretation suggests that younger children are more sensitive to and more affected by the particular processes whereby teachers communicate their expectations to children. 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.

A fifth interpretation suggests that the effects of teachers' expectations were more effective in the lower grade levels not because of any difference associated with the children's age but rather with some correlated sampling "errors." Thus it is possible that the children of the lower grades are the children of families that differ systematically from the families of the children of the higher grade levels.

A sixth interpretation also suggests that the greater IQ gain in younger children attributable to teacher expectation is a result of sampling "error," not in the sampling of children this time but in the sampling of teachers. It may be that in a variety of demographic, intellectual, and personality variables, the teachers of the younger children differed from the teachers of the older children such that they may have (1) believed the communications about their "special" children more or (2) been more effective communicators to their children of their expectations for the children's performance.

There is some evidence to suggest that teachers of the lower grades do in fact differ from the teachers of the upper grades of Oak School. Two administrators who were well acquainted with all the teachers rated them on over-all effectiveness as teachers. The two administrators agreed well in their ratings ($r = +.88$) and, although there were many exceptions, teachers of the lower grades were judged to be more effective teachers by both administrators (average r between effectiveness and teaching grade $= -.57$, $p < .02$).

The finding that only the younger children profited after one year from their teachers' favorable expectations helps us to understand better the results of two other experimenters, Clifford Pitt (1956) and Charles Flowers (1966). Pitt, it will be recalled, divided his sample of fifth-grade boys into three groups. For one group he reported the boys' IQ scores to the teachers after having arbitrarily added ten points. For another group he reported the boys' IQ scores after having deducted ten points. For the third group he reported the boys' actual IQ scores. Pitt found that there were no effects on school achievement at the end of the year of teachers having been given false information about their pupils' IQ.

The results of our own study suggest that after one year, fifth graders may not show the effects of teacher expectations though first and second graders do. Pitt's study differed in too many ways from our own to make direct comparisons possible, however. Pitt did not, for example, retest the children on IQ *per se* but only on school achievement. More important perhaps, is the fact that Pitt's teachers knew their pupils for nearly two months before being given pupils' IQ scores. That was long enough for teachers to have developed realistic expectations of pupils' performance more powerful than the expectations that could have been induced by adding or deducting IQ points.

The equivocal results of Flowers' experiment are also not directly comparable to our own data. Flowers' pupils were also older children (seventh graders) and each child had many different teachers rather than just one. Perhaps the effects of teachers' expectations were diluted by being distributed over many teachers. In the case of Flowers' study, we must bear in mind, too, that the classes arbitrarily labeled as brighter had been assigned different teachers than had been assigned to the control-group classes. Therefore, any differences between the experimental- and control-group classes could have been due to differences in the quality of teachers assigned to each. ■