In the mid-1950s, the results of my doctoral dissertation were nearly ruined; it appeared that I might have treated my experimental subjects in such a way as to lead them to respond in accordance with my experimental hypothesis, or expectancy.¹

To investigate the generality of these interpersonal expectancy effects in the laboratory, my colleagues and I conducted two studies employing animal subjects. Half the experimenters were told their rats were specially bred for poor maze (or Skinner box) performance, and half were told their rats had been specially bred for good maze performance. In both experiments, when experimenters had been led to expect better learning from their rat subjects, they obtained better learning from their rat subjects.²

At the end of the school year, 8 months later, all the children were retested with the same test of intelligence. Overall, the children from whose teachers had been led to expect greater intellectual gain showed a significantly greater gain than did experimenters expecting low ratings.³

The classroom study was in the minds of the teachers of these children. It was, then, in the minds of the teachers, that their scores on the "Test of Inflected Acquision" indicated they would show surprising gains in intellectual competence during the next 8 months of school. The only difference between the experimental group and the control group children, then, was in the minds of the teachers.

Unfortunately, evaluating the size of the effect that can be attributed to interpersonal expectancy is problematic because all three of the children in the study were included in the control group and the experimental group, thereby supporting the "Pygmalion" hypothesis.

To determine the practical importance of these results, measures of effect size are used. These measures are comparative values that express the magnitude of the difference between means in terms of a standardized mean difference. Such effect-size measures may be provided in one of several forms: the standard deviation of the difference between means, the standardized difference between means, the one or two-tailed binomial-probability, the one or two-tailed binomial-effect-size display (BESD). Table 1 shows the average magnitude of the effect obtained in the classroom study, which was .30, and the overall average magnitude of the effect obtained across the eight domains of research, with an overall size of .09.⁴

*The Harvard Test of Inflected Acquisition* indicated they were composed of children with above-average ability, average ability, and below-average ability, respectively. Within each of the 18 classrooms, approximately 20% of the children were chosen at random to form the experimental group. The

*Note:* The views expressed in this article are those of the authors and do not necessarily reflect the views of the United States Department of Defense or the Department of the Army.

Robert Rosenthal

1. Rosenthal, Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138.

2. Published by Cambridge University Press
attributable to the independent variable. To employ the BESD, one calculates the success rates for the treatment and control groups by adding \( \frac{1}{2}r \) to .50 (treatment group) and subtracting \( \frac{1}{2}r \) from .50 (control group). Table 2 illustrates the BESD for the overall mean effect size \( r = .30 \) of the 464 studies of interpersonal expectancy effects and, for comparison, for the results of a recent study of the effects of aspirin in the prevention of heart attacks \( r = .04 \).9

### Table 2. Two binomial-effect-size displays

<table>
<thead>
<tr>
<th>Condition</th>
<th>Favorable</th>
<th>Unfavorable</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>Pygmalion effects (( r = .30, d = 0.63, r^2 = .09 ))</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>Control</td>
<td>35</td>
<td>65</td>
<td>100</td>
</tr>
<tr>
<td>Treatment</td>
<td>Aspirin effects (( r = .04, d = 0.08, r^2 = .00 ))</td>
<td>52</td>
<td>48</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>52</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>200</td>
</tr>
</tbody>
</table>

Note. The Pygmalion effects are based on 464 studies. The aspirin effects are based on 22,000 participants.
Fig. 1. Ten-arrow model for the study of interpersonal expectancy effects.

Table 3. Four factors in the mediation of teacher expectancy effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Summary of the evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central factors:</td>
<td></td>
</tr>
<tr>
<td>1. Climate (affect)</td>
<td>Teachers appear to create a warmer socio-emotional climate for their “special” students. This warmth appears to be at least partially communicated by nonverbal cues.</td>
</tr>
<tr>
<td>2. Input (effort)</td>
<td>Teachers appear to teach more material and more difficult material to their “special” students.</td>
</tr>
<tr>
<td>Additional factors:</td>
<td></td>
</tr>
<tr>
<td>3. Output</td>
<td>Teachers appear to give their “special” students greater opportunities.</td>
</tr>
</tbody>
</table>

On the basis of the first 30 or so published studies relevant to mediation, a four-factor “theory” of the mediation of teacher expectancy effects was proposed. Table 3 summarizes these four factors, and Table 4 gives the average magnitude of the role of each factor separately for the B–C and C–D links. Although all four factors had significant effects, the magnitudes of the effects for the climate and input factors were especially impressive. Teachers appear to teach more and to teach it more warmly to students for whom they have more favorable expectations.

From these results, one cannot infer that selecting warmer teachers who present more material will result in children learning more. One also cannot infer from these results that training teachers to be warmer and to present more material will...

Published by Cambridge University Press
by judges who believe the defendants to be guilty, that depression among nursing home residents can be reduced by raising the expectations of caretakers, and that teachers' expectations can serve as self-fulfilling prophecies in other countries and for more than simply intellectual tasks. In all these cases, the mediating variables are receiving special attention, with the growing evidence indicating that much of the mediation is occurring by means of unintended nonverbal behavior.  

Acknowledgments—I thank the many students, colleagues, collaborators, and tutors who have been educating me for more than 40 years. Much of the research reviewed here was supported in part by the National Science Foundation, and preparation of this review was supported in part by the Spencer Foundation, though the contents are solely the responsibility of the author.

Notes
1. The details are given in R. Rosenthal, From over the Teacher Expectancy Effect, in the Fifth Sourcebook on Research Methods in Social and Behavioral Science, Washington, DC, D.P. Ray, Ed. (National Institute of Mental Health, 1961). That this research was received with ambivalence is illustrated by the receipt of two letters on the same day. The first letter rejected the paper for publication in a prestigious journal, and the second letter announced that the paper had received the Social Psychological Prize for 1960 from the American Association for the Advancement of Science.

2. R. Rosenthal and K.L. Fode, The problem of pygmalion in the nursing home: The effect of teacher expectancy on student IQ in a nursing home population, in The Behavior of the Young Child: Instrumental and Emotional Growth, pl. M. B. Balat, Ed. (Praeger, New York, 1968). A surprising finding was that the more children in the control group gained in IQ, the more unfavorably they were judged by their teachers. Apparently there were hazards to unpredicted intellectual growth. Also surprising was the strength of both the favorable and the unfavorable reactions to our research. For a summary of the criticisms and replies to them, see R. Rosenthal, Pygmalion effects: Existence, magnitude, and social importance, Educational Researcher, 16, 37–41 (1987).

3. Effect sizes are expressed in terms of both d and r. The former is the difference between the experimental and control groups divided by the standard deviation of both groups combined. The latter is the point-biserial correlation between experimental versus control group status (e.g., coding 1 for experimental and 0 for control) and the outcome score (e.g., gain in performance). The effect sizes in Table 1 are based on R. Rosenthal and D.B. Rubin, Interpersonal expectancy effects: The first 350 studies, The Behavioral and Brain Sciences, 3, 377–386 (1978).


5. Effect sizes are expressed in terms of both d and r. The former is the difference between the experimental and control groups divided by the standard deviation of both groups combined. The latter is the point-biserial correlation between experimental versus control group status (e.g., coding 1 for experimental and 0 for control) and the outcome score (e.g., gain in performance). The effect sizes in Table 1 are based on R. Rosenthal and D.B. Rubin, Interpersonal expectancy effects: The first 350 studies, The Behavioral and Brain Sciences, 3, 377–386 (1978).


8. The contents are solely the responsibility of the author.

9. For meta-analyses making this distinction, see M.J. Harris and R. Rosenthal, Mediation of interpersonal expectancy effects: 31 meta-analyses, Psychological Bulletin, 97, 363–386 (1985); M.J. Harris and R. Rosenthal, Four factors in the mediation of teacher expectancy effects, in The Social Psychology of Education, R.S. Feldman, Ed. (Cambridge University Press, New York, 1986). Although all the arrows in the model point to the right (i.e., from the past to the future), they may usefully be viewed as often going in both directions. Thus, improved student performance (D) can affect subsequent teacher expectations (B) and behavior toward the student (C).

10. R. Rosenthal, Pygmalion in Management and Their Mediating Mechanisms (MSS Modular Social Systems, Inc., 1985). A surprising finding was that the more children in the control group gained in IQ, the more unfavorably they were judged by their teachers. Apparently there were hazards to unpredicted intellectual growth. Also surprising was the strength of both the favorable and the unfavorable reactions to our research. For a summary of the criticisms and replies to them, see R. Rosenthal, Pygmalion effects: Existence, magnitude, and social importance, Educational Researcher, 16, 37–41 (1987).

11. For meta-analyses making this distinction, see M.J. Harris and R. Rosenthal, Mediation of interpersonal expectancy effects: 31 meta-analyses, Psychological Bulletin, 97, 363–386 (1985); M.J. Harris and R. Rosenthal, Four factors in the mediation of teacher expectancy effects, in The Social Psychology of Education, R.S. Feldman, Ed. (Cambridge University Press, New York, 1986). Although all the arrows in the model point to the right (i.e., from the past to the future), they may usefully be viewed as often going in both directions. Thus, improved student performance (D) can affect subsequent teacher expectations (B) and behavior toward the student (C).

12. R. Rosenthal, Pygmalion in Management and Their Mediating Mechanisms (MSS Modular Social Systems, Inc., 1985). A surprising finding was that the more children in the control group gained in IQ, the more unfavorably they were judged by their teachers. Apparently there were hazards to unpredicted intellectual growth. Also surprising was the strength of both the favorable and the unfavorable reactions to our research. For a summary of the criticisms and replies to them, see R. Rosenthal, Pygmalion effects: Existence, magnitude, and social importance, Educational Researcher, 16, 37–41 (1987).

13. D. Eden, Pygmalion's PONS: Some models for the study of interpersonal expectancy effects, in The Clever Hans Phenomenon, Annals of the New York Academy of Sciences, Vol. 364, T.A. Sebeok and R. Rosenthal, Eds. (New York Academy of Sciences, New York, 1981). A surprising finding was that the more children in the control group gained in IQ, the more unfavorably they were judged by their teachers. Apparently there were hazards to unpredicted intellectual growth. Also surprising was the strength of both the favorable and the unfavorable reactions to our research. For a summary of the criticisms and replies to them, see R. Rosenthal, Pygmalion effects: Existence, magnitude, and social importance, Educational Researcher, 16, 37–41 (1987).

14. R. Rosenthal and D.B. Rubin, A simple general-purpose display of magnitude of experimental effect, Journal of Educational Psychology, 74, 53–55 (1980). Both Raudenbush and Smith found effect sizes (r = .04, r = .04) for control) and the outcome of experimental and control groups divided by the standard deviation of both groups combined. The latter is the point-biserial correlation between experimental versus control group status (e.g., coding 1 for experimental and 0 for control) and the outcome score (e.g., gain in performance). The effect sizes in Table 1 are based on R. Rosenthal and D.B. Rubin, Interpersonal expectancy effects: The first 350 studies, The Behavioral and Brain Sciences, 3, 377–386 (1978).

15. L.A. Learman, J. Avorn, D.E. Everitt, and R.-S. Feldman, Eds. (Cambridge University Press, New York, 1986). Although all the arrows in the model point to the right (i.e., from the past to the future), they may usefully be viewed as often going in both directions. Thus, improved student performance (D) can affect subsequent teacher expectations (B) and behavior toward the student (C).


This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.