

A 30-Year Perspective

own laboratory, and with several ex-

puzzling and disconcerting re-

experimenter bias" that had led to

terpersonal expectancy effects in the

question about the possibility of in-

course, but it did raise a sobering

Al! of this was quite unwitting, of

respond in accordance with my exper-

such a way as to lead them to re-

trated my experimental subjects in

ducted with human subjects. Exper-

perimenters rather than just one.

sumably my students and I could

results of my dissertation, then pre-

pectancy effect or my "unconscious

psychological laboratory. If it was

ruined; it appeared that I might have

my doctoral dissertation were nearly

Social Psychology at Harvard

is Professor of

Robert Rosenthal

Pygmalion Effects in

PYGMALION EFFECTS IN

classrooms, approximately 20% of

were composed of children with

grade level, the three classrooms

rooms in the school, three at each of

"The Harvard Test of Inflected Ac-

"blooming." The test was labeled

administered a nonverbal test of intel-

brighter when expected to by their

menters becoming brighter when ex-

experimenters expecting low ratings.'^ 

half the experimenters led to expect

high ratings obtained substantially

percent of the children of the experimen-

ters had been led to expect bet-

menters had been led to expect bet-

maze (or Skinner box) performance, in

maze (or Skinner box) performance,

had been specially bred for good

had been specially bred for poor maze

presented with the same test of intel-

months later, all the children were

months of school. The only differ-

ence between the experimental

group and the control group chil-

children from

The earliest studies were con-

iments in the laboratory, my col-

these interpersonal expectancy ef-

earlier studies, experimenters expect-

half led to expect

photographs of people, with

rate photographs of people, with

rat subjects.\''

rat subjects.\''

ter learning from their rat subjects,

ter learning from their rat subjects,

think that children could become

think that children could become

states that their scores on the "Test of In-

states that their scores on the "Test of In-

nelence, which was disguised as a

nelence, which was disguised as a

that children would show surprising gains in intel-

that children would show surprising gains in intel-

nscious that experimental subject would show sur-

nscious that experimental subject would show sur-

"The Harvard Test of Inflected Ac-

"The Harvard Test of Inflected Ac-

tentation of the effect of a behavioral or

tention of the effect of a behavioral or

from the same problem: the under-

from the same problem: the under-

Table 1 gives an example of the

Table 1 gives an example of the

above-average ability, average ability, and below-

above-average ability, average ability, and below-

The Harvard Test of Inflected Ac-

The Harvard Test of Inflected Ac-

general estimate of the effect-size estimates suffer

general estimate of the effect-size estimates suffer

havioral intervention. Something

havioral intervention. Something

ences, with an overall

ences, with an overall

464 studies, with an overall

464 studies, with an overall

A dozen years after the Pygma-

A dozen years after the Pygma-

The earliest studies were con-

The earliest studies were con-

magnitudes of the effect ob-

magnitudes of the effect ob-

search. For each of these domains.

search. For each of these domains.

 unusual because all three of

 unusual because all three of

magnitudes of the effect ob-

magnitudes of the effect ob-

search. For each of these domains.

search. For each of these domains.

A dozen years after the Pygma-

A dozen years after the Pygma-

magnitudes of the effect ob-

magnitudes of the effect ob-

search. For each of these domains.

search. For each of these domains.

magnitudes of the effect ob-

magnitudes of the effect ob-

search. For each of these domains.

search. For each of these domains.

magnitudes of the effect ob-

magnitudes of the effect ob-

search. For each of these domains.

search. For each of these domains.

magnitudes of the effect ob-

magnitudes of the effect ob-

search. For each of these domains.

search. For each of these domains.
Table 1. Interpersonal expectancy effects in eight research domains

<table>
<thead>
<tr>
<th>Domain</th>
<th>Weighted Mean</th>
<th>Unweighted Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory interviews</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time</td>
<td>0.70</td>
<td>0.74</td>
<td>0.70</td>
</tr>
<tr>
<td>Learning and ability</td>
<td>0.54</td>
<td>0.55</td>
<td>0.84</td>
</tr>
<tr>
<td>Person perception</td>
<td>1.05</td>
<td>1.73</td>
<td></td>
</tr>
<tr>
<td>Inkblot tests</td>
<td>0.88</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>Everyday situations</td>
<td>0.70</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>Psychophysical judgments</td>
<td>0.70</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Animal learning</td>
<td>0.70</td>
<td>0.33</td>
<td></td>
</tr>
</tbody>
</table>

Note. The effect sizes shown are based on Rosenthal and Rubin. Weighting is by number of studies in each domain.

Table 2. Two binomial-effect-size displays

<table>
<thead>
<tr>
<th>Condition</th>
<th>Treatment</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favorable</td>
<td>Pygmalion</td>
<td>.30</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Aspirin</td>
<td>.04</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Unfavorable</td>
<td>Pygmalion</td>
<td>.09</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Aspirin</td>
<td>0.08</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>83</td>
</tr>
</tbody>
</table>

Note. The Pygmalion effects are based on 464 studies. The aspirin effects are based on 22,000 participants.

The 10-arrow model, designed to clarify the study of interpersonal expectancy effects, posits 10 links between five groups of variables: (a) distal independent (moderator) variables (e.g., stable attributes of the expecter and expectee), (b) proximal independent variables (the expectancies), (c) mediating variables, (d) proximal dependent variables (e.g., short-term outcome measures such as achievement on tests), and (e) distal dependent variables (long-term outcome variables). A useful feature of this model is that its 10 arrows represent the types of relationships that can be examined in research on expectancy effects (see Fig. 1). The arrows in the model are described in detail elsewhere, so I discuss here only the two links relevant to the topic of mediation: the B-C and C-D arrows. B-C relationships describe the effect of the expectancy on the expecter's behavior. These are the relationships most often investigated in research on mediation. Equally important to understanding mediation, however, are the C-D relationships between the expecter's behavior and outcome variables. Research bearing on the B-C link tells which behaviors are induced by a given expectancy, but research bearing on the C-D link shows that these behaviors affect the expectee so as to create a self-fulfilling prophecy. As is evident, the two types of relationships address different questions, making the distinction between them critical.
THE FOUR-FACTOR THEORY

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 lead to improved learning. The results do suggest, however, that conducting the research required to determine the benefits of selection and training for climate (or affect) and input (or effort) may well yield substantial benefits both for science and for society.

Looking forward, the most recent work has examined interpersonal expectancy effects in an ever-widening circle of contexts. Pygmalion effects in management, in courtrooms, in nursing homes, and in a variety of classrooms are under investigation. Studies have demonstrated that organizational effectiveness can be increased by raising leaders' expectations, that verdicts of guilty can be increased by having juries instructed...
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 fulfilling prophecies in other countries. 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.17

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 the contents are solely the responsibility of the author.

Notes


2. R. Rosenthal and K.L. Fode, The problem of experimenter-outcome bias, in Series Research in Social Psychology, D.P. Ray, Ed. (National Institute of Social and Behavioral Science, Washington, DC, 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 Sergei Psychological Prize for 1963 from the American Association for the Advancement of Science.


4. R. Rosenthal and L. Jacobson, Pygmalion in the Classroom (Holt, Rinehart and Winston, 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).

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 345 studies, The Behavioral and Brain Sciences, 3, 377–386 (1978).


7. Meta-analyzed data for the effects of experimenter bias on teachers' expectations of pupil IQ test performance are also available: S.W. Raudenbush, Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments, Journal of Educational Psychology, 76, 85–97 (1984); M.L. Smith, Teacher expectations, Evaluation in Education, 4, 53–55 (1980). Both Raudenbush and Smith found significant overall effects of interpersonal expectancy effects (r = .67) that substantial effects of teachers' expectations could be demonstrated only when the inducement was credible (i.e., when teachers had known pupils only 2 weeks or less at the time they were given the expectation). For the seven studies in which teachers had known pupils 1 week or less, as was the case for the Pygmalion study, the mean effect-size d was 0.29, and r was .14.


12. 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).


Recommended Reading


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.