

Interpersonal Expectancy Effects: A 30-Year Perspective

Robert Rosenthal

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 All of this was quite unwitting, of course, but it did raise a sobering question about the possibility of interpersonal expectancy effects in the psychological laboratory. If it was my unintentional interpersonal expectancy effect or my "unconscious experimenter bias" that had led to the puzzling and disconcerting results of my dissertation, then presumably my students and I could produce the phenomenon in our own laboratory, and with several experimenters rather than just one.

EARLY LABORATORY EXPERIMENTS

The earliest studies were conducted with human subjects. Experimenters instructed the subjects to rate photographs of people, with half the experimenters led to expect high ratings and half led to expect low ratings. In the first several such studies, experimenters expecting high ratings obtained substantially

Robert Rosenthal is Professor of Social Psychology at Harvard University, where he studies interpersonal expectancy effects, non-verbal communication, and data-analytic procedures. Address correspondence to Robert Rosenthal, Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138.

higher ratings than did experimenters expecting low ratings.²

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 had been specially bred for good maze (or Skinner box) performance, and half were told their rats had been specially bred for poor maze (or Skinner box) 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.³

PYGMALION EFFECTS IN THE CLASSROOM

If rats became brighter when expected to by their experimenters, then it seemed not farfetched to think that children could become brighter when expected to by their teachers. Jacobson and I conducted a study to test this hypothesis. All of the children in the study were administered a nonverbal test of intelligence, which was disguised as a test that would predict intellectual "blooming." The test was labeled "The Harvard Test of Inflected Acquisition." There were 18 classrooms in the school, three at each of the six grade levels. Within each grade level, the three classrooms 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

teachers of these children were told that their scores on the "Test of Inflected Acquisition" 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.

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 whom the teachers had been led to expect greater intellectual gain showed a significantly greater gain than did the children of the control group, thereby supporting the "Pygmalion" hypothesis.⁴

GENERILITY AND MAGNITUDE OF THE EFFECT

A dozen years after the Pygmalion-in-the-classroom study was completed, the research literature on interpersonal expectancy effects had broadened to include 345 experiments in eight domains of research. For each of these domains, Table 1 gives an example of the types of studies conducted and the average magnitude of the effect obtained.⁵ After another dozen years had gone by, there were 464 studies, with an overall *d* of 0.63, *r* of .30, and *r*² of .09.^{6,7}

Unfortunately, evaluating the social importance of these results is problematic because all three of these effect-size estimates suffer from the same problem: the underestimation of the practical importance of the effect of a behavioral or biomedical intervention. Something more is needed to address the question: What is the effect on the success rate (e.g., survival rate, cure rate, improvement rate) of any intervention? One approach is to use the binomial-effect-size display (BESD), which shows the change in outcome

Table 1. *Interpersonal expectancy effects in eight research domains*

Domain	Mean effect size		Example of type of study
	<i>d</i>	<i>r</i>	
Laboratory interviews	0.14	.07	Effects of sensory restriction on reports of hallucinatory experiences
Reaction time	0.17	.08	Latency of word associations to certain stimulus words
Learning and ability	0.54	.26	IQ test scores, verbal conditioning (learning)
Person perception	0.55	.27	Perception of other people's success
Inkblot tests	0.84	.39	Ratio of animal to human Rorschach responses
Everyday situations	0.88	.40	Symbol learning, athletic performance
Psychophysical judgments	1.05	.46	Ability to discriminate tones
Animal learning	1.73	.65	Learning in mazes and Skinner boxes
Weighted mean ^a	0.70	.33	
Unweighted mean	0.74	.35	
Median	0.70	.33	

Note. The effect sizes shown are based on Rosenthal and Rubin.⁵

^aWeighting is by number of studies in each domain.

attributable to the independent variable.⁸ 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

THE 10-ARROW MODEL

For many years, the central question in the study of interpersonal expectancy effects was whether they existed. The meta-analytic evidence

has answered that question sufficiently so that simple replications will add little new knowledge. Today, the central focus in the study of interpersonal expectancy effects has changed to include the isolation of the variables that (a) moderate expectancy effects and (b) mediate expectancy effects. Moderator variables are preexisting variables, such as sex, age, and personality, that influence the magnitude of interpersonal expectancy effects; mediating variables are the behaviors by which expectations are communicated.

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,¹⁰ 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 dis-

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

Condition	Treatment result		Total
	Favorable	Unfavorable	
Pygmalion effects ($r = .30, d = 0.63, r^2 = .09$)			
Treatment	65	35	100
Control	35	65	100
Total	100	100	200
Aspirin effects ($r = .04, d = 0.08, r^2 = .00$)			
Treatment	52	48	100
Control	48	52	100
Total	100	100	200

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

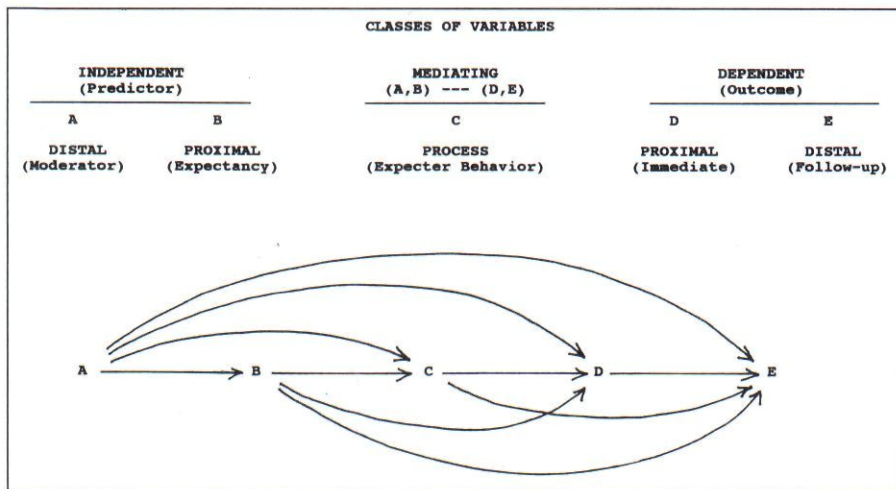


Fig. 1. Ten-arrow model for the study of interpersonal expectancy effects.

inction between B–C and C–D critical.¹¹

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.¹² 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.

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

Factor	Summary of the evidence
Central factors:	
1. Climate (affect)	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.
2. Input (effort)	Teachers appear to teach more material and more difficult material to their "special" students.
Additional factors:	
3. Output	Teachers appear to give their "special" students greater opportunities for responding. These opportunities are offered both verbally and nonverbally (e.g., giving a student more time to answer a question).
4. Feedback	Teachers appear to give their "special" students more informative feedback, both verbal and nonverbal, as to how these students have been performing.

Table 4. Meta-analytically derived average correlations representing the importance (effect sizes) of the four factors

Factor	Correlation between expectation and expecter's behavior (B–C link)	Correlation between behavior of expecter and response of expectee (C–D link)	Geometric mean correlation
1. Climate (affect)	.23	.36	.29
2. Input (effort)	.26	.28	.27
3. Output	.18	.16	.17
4. Feedback	.13	.08	.10

Note. These data are from Harris and Rosenthal (1986).¹¹ All correlations are significantly greater than zero at $p < .002$. The correlation between the magnitudes of the average B–C and C–D links is .88.

LOOKING FORWARD

The most recent work has examined interpersonal expectancy effects in an ever-widening circle of contexts. Pygmalion effects in management,¹³ in courtrooms,¹⁴ in nursing homes,¹⁵ and in a variety of classrooms¹⁶ 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 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.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 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 unconscious experimenter bias to teacher expectancy effects, in *Teacher Expectancies*, J.G. Dusek, V.C. Hall, and W.J. Meyer, Eds. (Erlbaum, Hillsdale, NJ, 1985).
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 Socio-Psychological Prize for 1960 from the American Association for the Advancement of Science.
3. R. Rosenthal and R. Lawson, A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats, *Journal of Psychiatric Research*, 2, 61–72 (1964).

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).

6. These meta-analyses are also described in R. Rosenthal, Teacher expectancy effects: A brief update 25 years after the Pygmalion experiment, *Journal of Research in Education*, 1, 3–12 (1991). For technical details on these meta-analytic procedures, see R. Rosenthal, *Meta-Analytic Procedures for Social Research*, rev. ed. (Sage Publications, Newbury Park, CA, 1991).

7. More specialized meta-analyses for the effects on pupils' 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 expectations on students' IQ. Smith reported a d of 0.16 and r of .08. Raudenbush found very strong evidence ($r = .67$) that substantial effects of teachers' expectancies could be demonstrated only when the induction of the expectancy 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.

8. R. Rosenthal and D.B. Rubin, A simple general purpose display of magnitude of experimental effect, *Journal of Educational Psychology*, 74, 166–169 (1982).

9. Steering Committee of the Physicians Health Study Research Group, Preliminary report: Findings from the aspirin component of the ongoing physi-

cians' health study, *The New England Journal of Medicine*, 318, 262–264 (1988). This study was discontinued on the grounds that it would be unethical to continue depriving placebo-control participants of the beneficial effects of aspirin, given the effect size ($r = .04$, $r^2 = .00$).

10. R. Rosenthal, Pavlov's mice, Pfungst's horse, and 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); also see Rosenthal, note 1.

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 (or 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, The mediation of Pygmalion effects: A four factor "theory," *Papua New Guinea Journal of Education*, 9, 1–12 (1973); R. Rosenthal, *On the Social Psychology of the Self-Fulfilling Prophecy: Further Evidence for Pygmalion Effects and Their Mediating Mechanisms* (MSS Modular Publications, New York, 1974).

13. D. Eden, *Pygmalion in Management* (Lexington Books, Lexington, MA, 1990).

14. P.D. Blanck, R. Rosenthal, A.J. Hart, and F. Bernieri, The measure of the judge: An empirically-based framework for exploring trial judges' behavior, *Iowa Law Review*, 75, 653–684 (1990).

15. L.A. Learman, J. Avorn, D.E. Everitt, and R. Rosenthal, Pygmalion in the nursing home: The effects of caregiver expectations on patient outcomes, *Journal of the American Geriatrics Society*, 38, 797–803 (1990).

16. E. Babad, Teacher expectancies and nonverbal behavior, in *Applications of Nonverbal Behavioral Theories and Research*, R.S. Feldman, Ed. (Erlbaum, Hillsdale, NJ, 1992).

17. N. Ambady and R. Rosenthal, Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis, *Psychological Bulletin*, 111, 256–274 (1992).

Recommended Reading

- Rosenthal, R. (1976). *Experimenter Effects in Behavioral Research* (Irvington, New York).
- Rosenthal, R., and Jacobson, L. (1992). *Pygmalion in the Classroom* (Irvington, New York).

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.