

The Hawthorne defect: Persistence of a flawed theory

“Like other hallowed but unproven concepts in psychology, the so-called Hawthorne effect has a life of its own.”

By Berkeley Rice

Most students of social psych are familiar with, or had better be if they want to pass. For decades, countless textbooks, Ph.D. theses, journal articles, and learned panels have cited it as a possible explanation for everything from why juvenile criminals in experimental program decide to go straight to why insomniacs sleep better in the laboratory. Whenever psychologists gather, one I apt to hear mention of the Hawthorne effect-even though, as it happens, the effect was never actually demonstrated by the original study.

Proponents of the Hawthorne effect say that people who are singled out for a study of any kind may improve their performance or behavior not because of any specific condition being tested, but simply because of all the attention they receive.

Those who mention the effect usually want to cast doubt on whether a given social innovation, instructional method, or therapy is really responsible for the change in behavior.

Though the Hawthorne effect has been generalized to every kind of psychological study, it grew out of a pioneering series of experiments that tested the impact of improved working conditions on productivity. In typical accounts of the findings, current textbooks report:

“To the surprise of the researchers, every innovation had the effect of increasing productivity.” (Lawrence Wrightsman, *Social Psychology*, 3rd ed., Brooks/Cole.)

“Almost no matter what experimental conditions were imposed, increases in output occurred....The investigators had obviously influenced the subjects’ behavior merely by studying that behavior, and this phenomenon has become known as the Hawthorne effect.” (Kelly Shaver, *Principles of Social Psychology*, 2nd ed., Winthrop, 1981

The research at Western Electric’s Hawthorne plant, near Chicago, from 1924 to 1932, helped to launch a whole new approach to human relations in industry, an approach that underlies current attempts by American industry to motivate workers and increase productivity by redesigning job conditions.

Recent re-evaluations have raised serious doubt about the productivity increases achieved in the Hawthorne studies, and also about what caused them. The longstanding impression that the employees at Hawthorne were happy about the changes in their jobs, and enthusiastically cooperated with the company’s experiments, have also come under sharp attack. A controversial article in *American Psychologist* last August, written by two radical social psychologists, charges that the Hawthorne effect is simply the result of “capitalist bias” among modern industrial psychologists.”

Like a number of other once widely held but faulty theories in psychology, such as the belief in a racial basis for intelligence, the Hawthorne effect has a life of its own that seems to defy attempts to correct the record. The story of this myth’s growth and its recent debunking contains a moral of caution for behavioral researchers and those who uncritically accept their pronouncements.

The Hawthorne experiments were carried out by a team of production engineers at the plant, where Western Electric manufactured telephone equipment (and still does). Most accounts of the research concentrated on a single study involving a small group of telephone relay assembly workers. In fact, there were several other productivity studies at Hawthorne during that period, the results of which have been pretty much misinterpreted or ignored for 50 years. Those results conflict with, or at least fail to support, the notion of the Hawthorne effect.

The first three experiments at the Hawthorne plant tested the effects of different degrees of illumination on the work of groups of women who inspected parts, assembled relays, or wound coils of wire. According to later accounts, the women in each of these experiments worked progressively faster,

regardless of increases or decreases in lighting. The only account of these experiments published at the time, however, was a brief summary in an engineering newsletter. No detailed or documented report ever came out, and the original research data somehow disappeared.

In another experiment, the “Mica Splitting Test,” the researchers first monitored the output of five experienced women at their regular department workstations, where they split, measured, and trimmed mica chips used for insulation. Then they moved the women to a special test room where, unlike their cohorts, they received 10-minute rest breaks at 9:30 a.m., and 2:30 p.m. After a brief decline in performance following the move, the women’s output increased by an average of .15 percent and remained at that level for the duration of the experiment. When they returned to their department, losing the rest periods, their output dropped back to the original rate. Since no other conditions had changed, the researchers attributed the increase in output to the beneficial effects of rest periods, not to the effect of the experiment itself.

In still another study, the researchers observed the performance of a team of 14 men who assembled telephone terminals in the “bank wiring” department. They then moved these men to a special test room, without introducing any other changes in work or pay conditions. Despite the move to a separate experimental setting, the men’s output did not increase, as it would have if the Hawthorne effect had occurred. Throughout the one-year experiment, the group maintained a steady rate of two terminals per day. By means of an informal understanding, individual members adjusted their work rates to keep up with or wait for the rest of the group. Group disapproval slowed down those whom the team members branded as “slaves” or “speed kings.” If the team got ahead of its normal rate, during the morning, the members tended to ease up during the afternoon. Thus, there was no increase in output to be explained by any Hawthorne effect.

In their popular accounts of the Hawthorne studies, published during the 1930s, Elton Mayo and his protégé Leta Stetter Stetter, industrial psychologists at Harvard, and William Dickson, A Western Electric engineer, tended to ignore or discount the results of these other experiments. So did most subsequent authors—who based their work on May, Roethlisberger, and Dickson, rather than re-examine the original data.

Nearly everyone concentrated on the longest experiment at Hawthorne (April 1927-June 1932), involving only five telephone relay assemblers. In their regular department workstations, these young women assembled the relays from about 35 separate small parts, a process requiring modest skills of memory, dexterity, and hand-eye coordination. After monitoring their performance for two weeks without their knowledge, the researchers moved the five women to a separate relay-assembly test room in order to measure the effect of two variables: rest periods and the length of the workday.

As summarized in a currently popular textbook, this is what supposed to have happened: “Regardless of the conditions, whether there were more or fewer rest periods, longer or shorter workdays...the women worked harder and more efficiently. Although this effect was probably due to several reasons, the most important was that the women felt they were something special...that they were expected to perform exceptionally. They were happy, a lot of attention was paid to them, and they complied with what they thought the experimenter (their boss) wanted.” (Jonathan Freedman, et al, *Social Psychology*, 4th ed., Prentice Hall, 1981).

Enter H. McIlvaine Parsons, a distinguished industrial psychologist, past president of the Society of Engineering Psychologists and of the Human Factors Society. Like most of his colleagues, Parsons had accepted and often cited the Hawthorne effect without questioning the evidence. (Indeed, he says that it’s not even clear who first identified the supposed effect.) Prompted by a student’s query, however, his curiosity led him in the early 1970s to a lengthy re-examination of second- and then first-hand accounts of the research.

Many of those originally involved in the research had long since died, and few of those who had written about it later had actually observed the experiments themselves. When he heard that Western Electric was planning to celebrate the 50th anniversary of the experiments at the original red brick plant in

1974, Parsons got himself invited. There he had the good fortune to meet and interview Donald Chipman, one of the supervisor-observers in the relay test room experiment, and Theresa Zajac, one of the five experimental subjects. Nearly 70, she was still toiling faithfully at the same plant, 50 years later.

From his interviews and his analysis of the original research data, Parsons discovered not only serious gaps and flaws in the published reports of the Hawthorne experiments, but also a number of what he calls “confounding variables” that previous researchers had ignored. For example, unlike the big open floor of the relay-assembly department, the test room was separate, smaller, and quieter, with better lighting and ventilation. And the supervisors were friendly, tolerant observers, not the usual authoritarian foremen. Any or all of these factors may have contributed to the improved performance. The friendlier atmosphere of the test room also led to the most serious flaw in the experiment. After about three months, talking among the workers increased to the point that the experimenters feared that it would jeopardize the research. According to the observer’s log, four of the women were reprimanded and told that they would be taken back to the regular department unless their performance improved. After eight months of the experiment, according to Parsons, two of the women “persisted in talking too much while they worked. Their output rates were dropping, and therefore the group’s rate dropped.” The other women resented the resulting loss in their pay. After repeated warnings, the two women were dropped from the study, and replaced by two other, more cooperative relay assemblers. One of the replacements proved so industrious and enthusiastic that she soon became the group’s unofficial leader, often spurring her cohorts to work faster.

The replacement of two out of five subjects in mid-experiment, particularly for being too slow, seriously contaminates any subsequent findings of increased output by the group. But even the claims of continuing gains in productivity do not bear up under careful examination of the data. Actually, the group’s hourly output fell off noticeably—as one might have expected—during periods when the rest breaks were withdrawn and when the workday was lengthened. Over the course of the entire study, however, the group’s total output did rise by nearly 30 percent.

Parsons uncovered several possible reasons for the increases in output, in addition to the replacement of the two slow workers—reasons that seem far more plausible than the mysterious Hawthorne effect. For example, the researchers never seemed to consider the possibility that over the course of the five years, the women were simply becoming more skillful at their work—and thereby faster. As evidence, Parsons noted that at various times during the experiment, when the supervisor asked them to try to work as fast as they could for short intervals, he found that their rate of output rose considerably, showing the potential for improved performance. Since the women’s performance was not measured back at their old department, after leaving the test room, there’s no way of to judge whether such improvement occurred.

Parsons detected two important factors that anyone familiar with the theories of behavior modifications would spot, but that the Hawthorne researchers largely ignored. (Of course, these theories were not yet well known in the 1930s.)

Back in the relay-assembly department, the women had been paid a fixed hourly wage plus a collective piecework rate based on the department’s total output. In the test room, the collective piecework rate was based on the output of only the five workers, so that individual performance had a much more significant impact on weekly pay. The monetary reward for increased individual effort thus became much more evident and perhaps more effective than in the department setting.

This motivation effect was undoubtedly multiplied by another “confounding variable,” Parsons pointed out. Unlike the department shop floor, where relay assemblers had no regular or immediate access to individual production figures, the experimental test room provided plenty of what Parsons calls “performance feedback.” Separate counters recorded the number of relays assembled by each worker. The counters were accessible to any of the women who wanted to check them, and some did. The supervisor took regular readings from them and posted daily totals on the wall. The observer’s logs contain numerous comments by the women that show they were interested in, and kept close track of, their output. For example, on the afternoon of April 19, 1929, Theresa Zajac remarked “I’m about 15

relays behind yesterday.” Another woman said, “I made 421 yesterday, and I’m going to make better today.” As the observer commented, the women were “trying to beat their former output records.”

As evidence for the importance of performance feedback, Parsons cites the results of another relay-assembly experiment designed as a control group for the test-room study. Although the five women in the control group also worked under the small group pay rate, they worked at a separate bench on the department floor, and had no regular access to information about their individual daily output. And while their output also rose (by about 12 percent) after they switched from their regular pay schedule to the small-group rate, it remained at that level until they returned to their old work stations whereupon it dropped to the original level. There were no progressive increases as in the test-room experiment. In an article in *Science* (March 8, 1974), Parsons compared the two groups and concluded that “the faster the workers assembled relays, the more money they got, but knowledge of results was essential to make this differential reinforcement effective.”

Since subsequent research has failed to duplicate the supposed Hawthorne effect in various experimental settings, Parsons believes that it simply shows the effect of variables that experimenters are unaware of, or over which they have no control. In his writings and lectures, Parsons, now 70, has attempted to correct the mistaken interpretations of the Hawthorne studies, but he recognizes that the theory is so entrenched that it has become part of the accepted wisdom among social scientists. When asked why the authors of current textbooks continue to include unquestioningly the traditional account of the research, despite his well-publicized critique, Parsons replied, “They’re lazy.”

So much for the factual basis of the Hawthorne myth that productivity increased “regardless of the conditions tested.” Output, as Parsons showed, had not improved under all circumstances. As for the other part of the myth—that the workers’ performance improved because they knew they were part of an important experiment and responded enthusiastically to the attention paid to them—that notion has just been skewered by Daria Bramel and Ronald Friend of the State University of New York at Stony Brook, in their *American Psychologist* article.

Bramel and Friend’s analysis is heavily laden with Marxist ideology and a conspiratorial view of industrial psychologists as unwitting or even willing tools of capitalism, helping to manipulate and exploit the working class for the sake of corporate profits. Regardless of the accuracy of their view, their article reveals considerably evidence of suspicion, reluctance, and resistance among the workers involved in the Hawthorne experiments, which contrasts sharply with the “wholehearted cooperation” reported by Mayo and others.

According to one account cited by Bramel and Friend, the two talkative women in the relay test-room experiment were dropped for “gross insubordination” as well as low output. In a private letter, Elton Mayo referred to one of them as having “gone Bolshevik.” After interviewing her later, Roethlisberger reported: “She also said that she had heard comments from girls in the regular department to the effect that what the company really was after in the test room was maximum output, and that the test room was not being run, as the investigators said, to determine the best work conditions.”

As the experiment went on, with the introduction and then withdrawal of rest breaks and longer workdays, the women remained “somewhat suspicious and apprehensive,” engaging in frequent skirmishes with the observers and researchers. “Give me back the rests,” said one of the women, “and see how my output goes up.” Based on their reading of the original documents, Bramel and Friend interpreted the decreases in output not as an effect of the withdrawal of rest periods, but as a deliberate slowdown by the group designed to force their return.

Bramel and Friend argue that because of “elitist biases,” the psychologists who wrote about the Hawthorne experiments assumed that the discovery of more productive working conditions would somehow benefit the workers as well as management. Mayo, for example, referred to output as an “index of well-being” apparently assuming that a work group’s productivity corresponds to and depends on the morale of its members. (Recent research in industrial productivity does not support that assumption.) To factory owners and managers in the 1930s who followed the results of the Hawthorne studies, anything

that promised to increase output without raising pay scales seemed to be worth pursuing.

As the recent attempts to re-evaluate the Hawthorne experiment demonstrate, the studies continue to serve as a kind of Rorschach test for managers and industrial psychologists, enabling them to find evidence to support many different and often conflicting theories of how to motivate the modern industrial worker. (One recent analysis suggested that the increased output at Hawthorne was really caused by the deepening economic depression at the time, and the women's resulting need for more money.) This confusion may not be all bad. For whatever the flaws, in the conduct and subsequent interpretations of the Hawthorne studies, they did spur effort to humanize the workplace, to find more sensitive ways to mobilize workers, rather than regarding them as assembly line robots that could be kept producing by fear and discipline. The promise was that social engineering, supported by enlightened management and cooperative workers, could usher in a new era of industrial peace and prosperity. If this hope has since proven somewhat naïve, it was at least well intentioned.

-----Berkeley Rise is a senior editor of *Psychology Today*