Scand J Work Environ Health 2006;32(5):402-412

<https://doi.org/10.5271/sjweh.1036>

Issue date: 31 Oct 2006

The “Hawthorne effect” is a myth, but what keeps the story going?

by [Kompier MAJ](#)

Affiliation: Radboud University Nijmegen, Department of Work and Organizational Psychology, PO Box 9104, 6500 HE Nijmegen, Netherlands. M.Kompier@psych.ru.nl

Refers to the following text of the Journal: [2000;26\(4\):363-367](#)

Key terms: [discussion paper](#); [Hawthorne effect](#); [Hawthorne studies](#); [Human Relations School](#); [myth](#)

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/17091208

This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

The “Hawthorne effect” is a myth, but what keeps the story going?

by Michiel AJ Kompier, PhD

Kompier MAJ. The “Hawthorne effect” is a myth, but what keeps the story going? *Scand J Work Environ Health* 2006;32(5):402–412.

The Hawthorne studies became famous because of the discovery of the “Hawthorne effect”: “a marked increase in production related only to special social position and social treatment”. They mark the beginning of the Human Relations School. This article demonstrates that the Hawthorne research does not pass a methodological quality test. Even if methodological shortcomings were waived, there is no proof of a Hawthorne effect in the original data. The following five myths are debunked: (i) scientific worth, (ii) continuous improvement, (iii) social factors prevailing over physical factors and pay, (iv) wholehearted cooperation, and (v) the neurotic worker. The following five factors are held responsible for the creation and survival of the Hawthorne myth: (i) a story too good to be untrue, (ii) bias and selective accounts by original researchers and “laziness” among later scientists, (iii) social factors do matter, and (iv) a story that fits the cognitive world and interests of psychologists, and (v) management.

Key terms Hawthorne studies; Human Relations School.

The Hawthorne studies, conducted from 1924 to 1933 at the Western Electric Company’s Hawthorne plant, represent a major historical event in the development of social sciences. In industrial sociology or psychology, no other theory or set of experiments has stimulated more research and controversy nor contributed more to a change in management thinking than the Hawthorne studies and the human relations movement they spawned (1).

The research consisted of a series of experiments in order to study the effects of illumination, rest breaks, length of workday and workweek, wages, food, humidity, and temperature on worker performance. Moreover, more than 20 000 interviews with Hawthorne employees were conducted. These studies have become famous, not because of the light they shed on the nature of the aforementioned relations, but because of the “discovery” of a special effect, the “Hawthorne effect”. Textbooks characterize this effect more or less as follows: “No matter what the researchers did, productivity went up. Even when work conditions were made worse than they were originally, the women worked harder and more efficiently. The secret ingredient? The attention shown to them by all those concerned with the study

was the variable which influenced their behaviour” (2, p.372).

This conclusion has become part of the accepted wisdom among social scientists and intervention researchers. With time, it has become increasingly common to attribute any unexpected result occurring in an experiment with human participants to the Hawthorne effect (3). Accordingly, for generation after generation, millions of students have been raised with the story that people who are singled out for a study of any kind improve their performance or behavior not because of any specific condition being tested, but simply because of all of the attention they receive.

Historically, this conclusion marks the beginning of a new era in industrial psychology: the Human Relations School. This influential movement promotes worker-management harmony, while emphasizing the role of social factors, group processes, social skills, leadership, and “worker morale” in influencing work productivity.

The aim of this contribution is to (i) demonstrate that the Hawthorne effect is a persistent story; (ii) to analyze what actually happened in the Hawthorne research; (iii) to debunk five Hawthorne myths; and (iv) to analyze how this story could emerge and survive.

¹ Radboud University Nijmegen, Nijmegen, Netherlands.

Reprint requests to: Dr M Kompier, Radboud University Nijmegen, Department of Work and Organizational Psychology, PO Box 9104, 6500 HE Nijmegen, Netherlands. [E-mail: M.Kompier@psych.ru.nl]

Hawthorne effect: a persistent story

Primary sources (4–8) do not mention the term Hawthorne effect. The term was probably introduced by French (9), who described it as “a marked increase in production related only to special social position and social treatment [p 101]”. The appendix, which covers a 50-year time span, provides eight examples of definitions of the Hawthorne effect. It could easily be expanded with similar examples (10, 11).

What actually happened in the Hawthorne research?

The Western Electric Factory was the supplier of telephone equipment to the last American Bell System. The Hawthorne plant was its main factory. At the start of the studies, it offered employment to some 29 000 men and women, many of them immigrants. The Hawthorne research consisted of the following six, partly overlapping, studies: (i) three illumination experiments (November 1924–April 1927), (ii) the first relay assembly test room (April 1927–February 1933), (iii) the second relay assembly group (August 1928–March 1929), (iv) the mica splitting test room (October 1928–September 1930), (v) the interview program (September 1928–early 1931), and (vi) the bank wiring observation room study (November 1931–May 1932).

The researchers published their studies in five reports (4–8). The most thorough was that of Roethlisberger & Dickson (6), who presented data on all six studies. The Whitehead book (5) covers only, though in great detail, the first relay experiment in the assembly test room. Although Mayo, a Harvard business professor, was not the director of the studies [he first visited the factory in 1928 in the middle of the first relay experiment], he became the main interpreter and prophet of the Hawthorne studies. He wrote forewords for the reports of Whitehead (5) and Roethlisberger (7), and a preface for the publication by Roethlisberger & Dickson (6). Roethlisberger and Whitehead worked at Harvard, and Dickson was an officer of the Western Electric Company. By November 1932, the economic recession caused the suspension of the studies. Operators of the famous first relay assembly test room had all been laid off by then.

Illumination experiments, 1924–1927. The illumination studies are not always formally considered to be part of the Hawthorne studies. When, in the early 1920s, the electrical suppliers of Western Electric claimed that better lighting would improve productivity (12), the company started three illumination experiments. The National Research Council (NRC) sponsored the research. The data were never formally reported, and it is unknown how many employees were studied. The best available “official” description is a 4-page summary (out of 604

pages) in the book by Roethlisberger & Dickson (6, p 14–18). These authors repeatedly refer to a short tentative report by Snow (13), the NRC’s representative. The final report of this project was never delivered to the NRC (14, p 183). Snow concluded (experiment 1) that “The corresponding production efficiencies by no means followed the magnitude or trend of the lighting intensities. The output bobbed up and down without direct relation to the amount of illumination” (6, p 15). The researchers learned from these studies that light was not the only factor that influenced employee output and that the results could have been influenced by any one of several variables. They also decided that further studies should not be conducted in regular shop departments or on fairly large groups.

First relay assembly test room, April 1927–February 1933. The most famous experiment was the first study that took place in the relay assembly test room. Stimulated by the illumination studies, the researchers “decided to isolate a small group of workers in a separate room, somewhat removed from the regular workforce, where their behaviour could be studied carefully and systematically” (6, p 19). The experimental group consisted of five young women. A sixth “lay out operator” distributed the materials among these five operators (4, p 58) (figure 1). She sat next to operator 5, and her income was based on the output of the five operators. According to Whitehead (5, p 14) “The actual method of selection was quite informal and somewhat obscure”. The operators were selected because they were “thoroughly experienced” and “willing and cooperative”, and not to be married soon. After the first two operators were identified, these two invited the three other girls. The operators’ task was to assemble relays (electromagnetic switches) for telephones. A relay consisted of between 26 and 52 parts. The operators assembled over 150 different types between 1927 and 1932. Typical cycle times

Figure 1. The first relay assembly test group. [Source: Roethlisberger & Dickson (6), figure 3, p 24]

were within 1 minute, and an operator assembled approximately 500 relays each day.

In a very complicated order, depending on decisions of the experimenters and influenced by economic circumstances, two independent variables were manipulated, rest pauses and duration of work, in 24 experimental periods (and not 13 as often reported). In addition, in period 3 a piecework system was introduced. It was based on the average output of the experimental group and not, as before, on the output of the entire department. Changes were introduced cumulatively, and no control group was established. In the various periods (combinations of) rest pauses and work duration were introduced. For example, breaks were introduced in period 4, removed in period 12, and then re-installed. The number of workhours was decreased in periods 8 and 9 and restored in period 10 to the normal baseline level, and so forth.

In short, the researchers concluded that productivity increased clearly over time primarily because of the changed supervision of the workers. Rest pauses, shortening of the workweek, and the individualized reward system were considered less important. [See, for example, the reports of Mayo 4, p 65; 8, p 63–64, p 72]. Inspection of the original data reveals that there was no steady increase in output. For example, in period 12—no lunch and no rest pauses at all—hourly output rates clearly decreased for four of the five operators (6, figure 6, p 76; 5, p 88; and 5, figures B-4, volume 2). Operator 3 even said (6, p 72): "Give me back the rests and see how my output goes up".

The internal validity of this experiment is further invalidated by selective attrition. At the beginning of period 8 (January 1928) operators 1a and 2a were removed from the experimental group because they were too busy "talking and fooling". Note that the five operators were allowed to talk during work, whereas, in the regular department, the operators were expected to work in silence (5, p 111). Although they had been warned repeatedly and threatened with disciplinary action, they "did not display that wholehearted cooperation desired by the investigators" (6, p 53). "Their controlled experiment was being jeopardized, and something had to be done" (6, p 54). They were dismissed "for the best interest of the test" (p 55) because of gross insubordination and declining or static output. Whitehead speaks of the rejection of the obstructive minority (5, p 119), but Mayo tells the story differently, writing that operators 1a and 2a "dropped out" (4, p 58) and were "permitted to withdraw" (4, p 114–115). However, in a 1929 letter written to the Laura Spelman Rockefeller Memorial, Mayo wrote of operator 2a: "One girl, formerly in the test group, was reported to have 'gone Bolshevik' and had been dropped" (cited in 15, p 54; and 10, p 871). There is some evidence of this type of worker resistance

on the part of these two operators in Whitehead's report: "their attitude that we [the researchers] only wanted to make them establish a record and then forever after equal it" (5, p 118). Roethlisberger (7, p 13) denies these replacements: "During the first year and a half of the experiment, everybody was happy, both the investigators and the operators".

It is probable that new operator 2 had also taken part in the illumination studies (16). Whitehead (5, p 15) also reports that operator 5 resigned from the firm for family reasons in the middle of 1929 and rather unexpectedly returned 10 months later. In the meantime, a new substitute had been appointed, number 5a. The return of operator 5 caused trouble in the group because that would imply that 5a had to return to the regular shop where she would earn less (5, p 143). Most interestingly, new operators 1 and 2 (friends in the main shop) immediately proved to be high performers (6, figures 6, p 76). An analysis of the output of the five operators reveals that operator 2 "jumped immediately into the lead and maintained it throughout the experiment" (6, p 161–162). She also "held all records in speed tests and in hourly, daily, weekly and period output, and she received the highest scores in the various dexterity and intelligence tests" (6, p 167). The other new operator, operator 1, soon took second place. In the last two periods before her dismissal, operator 1a had ranked fifth. New operator 2 became the informal leader of the group. She pressed her colleagues to increase their output (5, p 123; 6, p 167).

Second relay assembly group, August 1928–March 1929.

The investigators wanted to single out the effect of the change in the method of payment on output and therefore started the next two studies. In the second relay study, five experienced relay assemblers (women) were selected by the foreman of the regular department and formed into a special group. They assembled relays under normal work conditions (eg, supervision) in a regular department, where they worked in adjacent positions at the same bench. Unlike their colleagues in the first experiment, they were not segregated into a separate room and remained in the regular department. However, whereas the normal workers' salary was based on the output of the entire department (approximately 100 employees), the experimental group received the same type of remuneration as the operators in the first relay assembly group: their salary was based on the average output of just the five women.

After a 5-week "base period", the experimental period followed. According to Roethlisberger & Dickson, average hourly output increased by 12.6% (6, table 12, p 132). Contrary to planning, the experimental period continued for only 9 weeks. After a few weeks, "the foreman began to report to the investigators that the

presence of a special group in his department was causing considerable friction among the other employees, as they too wanted similar consideration" (6, p 133). The problems quickly increased and "it became necessary, in order to preserve the department's morale, to return the operators to the regular method of payment" (6, p 133). The output of the five promptly dropped by 16% (third period: 7 weeks).

Many researchers would consider these rises and falls in outcomes as supporting evidence for the importance of financial incentives. Not Roethlisberger & Dickson, who repeatedly (6, p 158, p 577) and without proof suggested that the second group performed well because they wanted to equal the record of the first relay test group.

Mica splitting test room, October 1928–September 1930.

The objective of the study in the mica splitting test room was to investigate the effects of individual piecework (throughout the whole experiment) and overtime. Overtime implied a workweek of 55.5 hours. Occasionally the girls also worked on Sunday (and were then paid 100% extra). Furthermore, longer experimental periods were chosen. Apart from these experimental factors, the design of the mica experiment was as similar as possible to the first relay assembly test. Of the five women originally selected for this experiment, only two were willing to participate. They selected three other girls and formed a test group of five. Mica splitting was considered one of the most desirable Hawthorne-shop jobs for women. The work was highly repetitive, but required considerable skill (it took 2 to 3 years to really master it) and consequently was well paid.

Table 1 provides an overview of the five test periods. It shows that another experimental factor was manipulated, rests (two 10 minute-periods, at 0900 and at 1430) versus no rests. The last manipulation (period 5, a 40-hour workweek) was not deliberately chosen by the experimenters but the result of the recession.

During each of the five periods individual output was counted, and observers observed the operators and reported their opinions. A strong dislike for working on Sundays appeared. Operator M4 (1 March 1929): "I suppose we'll be working on Palm Sunday and Easter Sunday. We won't call it Sunday anymore, we'll call it Slave Day" (6, p 141). During the experiment the posi-

tion of the mica girls in the regular department was severely threatened. Much of the work was transferred to the Kearny plant in New Jersey. Job insecurity affected the five operators (6, p 143). Operator M4: "Now that the work is dropping off, every time the other girls [from the regular department] meet us, they ask, 'Are you still on mica?' When we tell them we are, I can see they are sorry although they don't say it". An overview of the average hourly output by week of the five operators is presented by Roethlisberger & Dickson 6 (p 147). On the basis of these data, the researchers concluded that output "tended to increase during the first year" (6, p 149). An inspection of the original data (6, p 147) indicates that this seems imprecise. First, when compared with the baseline period (period 1), there seems to have been a slight decline in average output when the girls were moved into the test room (period 2). Next, it appears that, when rest pauses were introduced in period 3, "a moderate but steady rise in rate of output" took place (6, p 146). The authors also pointed to a decline in output during the second year and attributed it to the considerable tension among the five, because they were afraid of losing their jobs (6, p 153).

Interview program, September 1928–early 1931. "As the test room studies had clearly indicated that there was a close relation between employee morale and supervision", the interview program started as a plan for improving supervision (6, p 189). The company wanted to determine what constituted "an effective working together of supervisors and employees" and decided to ask the employees themselves to express frankly their (dis)likes about their work environment. The first round of interviews took place in the Inspection Branch (approximately 1600 employees). Anonymity was guaranteed. The program was extended, first (in 1929) to the Operating Branch and eventually to all eight branches in the Hawthorne plant (1930). In these new rounds, the interview method was changed from a direct (resembling a questionnaire method) to an indirect approach (6, p 203; 7, p 20), because the researchers felt that they might miss what was really important to the employees if they themselves directed the interview. In the indirect approach that concentrated on the concept of "meaning", the employee was allowed "to choose his own topic". The average time for a single interview ac-

Table 1. Mica splitting test room—schedule of test periods. (Source: Roethlisberger & Dickson, 1939 (6), table XIV, p 138)

Period	Special feature	Dates included	Duration in weeks
I	In regular department; overtime, no rests	27 August 1928 to 20 October 1928	8
II	Moved to test room; overtime, no rests	22 October 1928 to 24 November 1928	5
III	Overtime with two 10-minute rests	26 November 1928 to 15 June 1929	29
IV	No overtime, two 10 minute rests	17 June 1929 to 17 May 1930	48
V	8-hour day, 5-day week, two 10-minute rests	19 May 1930 to 13 September 1930	17

cordingly increased from 30 to 90 minutes. All of the interviews were written down as near to verbatim possible, and an analyzing department, part of Hawthorne's special Industrial Research Division, analyzed the interviews. By 1930, a total of 21 126 employees had been interviewed. The interviews served as an input for supervisory training, for the improvement of unfavorable work conditions, and as research material. Table 2 presents the outcomes of the re-calculation of the original data (an overview of the 15 issues that received the most unfavorable comments in the 1929 study among 10 300 employees from the Operating Branch).

By far the highest number of negative comments concerned "payment" (rate revision: 1541, piecework rate: 1510, wages: 1284, piecework in general: 1123), followed by "lockers" and "safety and health". "Supervision" came fourth, followed by a long list of material work conditions. Roethlisberger & Dickson did not conclude that especially conditions of pay were considered important by the Hawthorne employees (6, p 240). From the interview program the researchers concluded that, apart from the formal Hawthorne organization, also several types of informal organization existed and that work groups controlled their output through social norms and corrective actions.

Bank wiring observation room study, November 1931–May 1932. In order to supplement the large interview program, an observational study was conducted to further study how work groups controlled their output through social norms and behavior. There were no experimental manipulations. Fourteen male operators were observed. Three solder men, two inspectors, and nine wiremen worked in a special room, on terminal banks for telephone exchanges. It was an already existing

Table 2. Fifteen issues that received the highest numbers of unfavorable comments in interviews in the operating branch (10 300 persons: 6800 men, 3500 women) in 1929. (Recalculated from Roethlisberger & Dickson, 1939 (6, p 232–5).

Issue	Number of unfavorable comments
1. Payment (eg, rate revision, piecework rate)	6816
2. Lockers	3540
3. Safety and health	3208
4. Supervision	2737
5. Workhours (eg, overtime)	2273
6. Washrooms	2044
7. Light	1689
8. Ventilation	1524
9. "Bogey" (a performance standard)	1384
10. Job placement	1318
11. Fatigue	1311
12. Dirt	1297
13. Tools and machines	1281
14. Material	1241
15. Temperature	1120

group that performed their regular duties and reported to their regular supervisors. As before, they were paid by department-wide piecework; thus their income was related to the output of approximately 100 workers. The men were observed by an observer, in "the role of an disinterested spectator" (6, p 388) and much like "an anthropological field worker" (6, p 389), and repeatedly interviewed.

The output of the group remained practically unchanged (6, p 424). The men practiced the concept of a "fair days' work": "the working group as a whole actually determined the output of individual workers by reference to a standard, pre-determined but never clearly stated, that represented the group conception of a fair day's work. This standard was rarely, if ever, in accord with the standards of the efficiency engineers" (8, p 70). If one operator worked faster, he was corrected by his co-workers who called him "slave", "rate buster", or "speed king". Another corrective practice was "binging", hitting him sharply on the upper arm. Operators that were too slow, "chiselers", also met with group disapproval. Group pressure for controlled output was typically explained by the men as follows (6, p 417): "If we exceed our day's work by any appreciable amount, something will happen. The 'rate' might be cut, the 'rate' might be raised, the 'bogey' [an output standard] might be raised, someone might be laid off, or the supervisor might 'bawl out' the slower men" (7, p 22).

Debunking five Hawthorne myths

Over time, these studies have been interpreted, reinterpreted, and often misinterpreted. Close examination of the original data as reported in the quoted primary sources leads to the conclusion that there are, in fact, the following five interrelated Hawthorne myths: (i) the myth of scientific worth, (ii) the myth of continuous improvement, (iii) the myth of social factors being more important than physical factors and pay, (iv) the myth of wholehearted cooperation, and (v) the myth of the neurotic worker.

The myth of scientific worth. An assessment of the methodological quality of these case studies (17) would result in the lowest qualification (ie, evidence that is descriptive, anecdotal or authoritative). These studies do not pass an elementary methodological quality test. There was a lack of scientific rigor, confounding of experimental factors, and so many uncontrolled variables that it became virtually impossible to identify any causal relationship (3, p 364). There was no control group (all studies). Utilizing a control group would have made it possible to compare the "experimental group outcomes" with the "natural flow of events". For example, in the mica splitting test room, it would have made sense to

compare the decline in output during the second year with the performance in a control or reference group. And a control group(s) design in the first study on the relay assembly test room could have made it less difficult to disentangle potential influences of ergonomic and social factors. Furthermore, there was a very small number of participants (all studies, except the interview program), and there was selection bias (motivated volunteers) and attrition that evidently had an impact on the outcomes (first relay assembly test). As to the illumination experiments, there are no data, and there is no official report.

The myth of continuous improvement. A main conclusion of the researchers was that “In the course of the test room experiment, the one outstanding factor which challenged interpretations was the general improvement in the output of the operators, which rose independently of the specific changes in conditions of work made during the study” (6, p 189). This conclusion was not justified in the first relay study (5, p 88), nor was it justified for the other experiments. In the first relay study hourly output rates clearly decreased for four of five operators after, in period 12, time for lunch and rest pauses were taken away. By complicated mixtures of the two measures “average hourly output” and “total weekly output”, this issue was obscured by the researchers. For example, Mayo’s discussion of period 12 of the relay assembly test room emphasized the increase in total output but initially ignored (4), and later significantly downplayed (8), the declining hourly output rate (see also 18, p 26). The account of Roethlisberger (7, p 13) is also not correct. He wrote that, in period 12, output “maintained its high level”. Moreover, Jones (19) utilized multivariate regression techniques to re-analyze the original data of the first relay study and “found essentially no evidence of Hawthorne effects” (“a common effect that could be regarded as a pure result of the experimentation”, p 467). In the bank wiring study the output of the group under study did not improve but remained at the same level. Clearly, in this case, being part of an experiment did not have an impact on the worker’s performance. There was also no steady increase in the mica test.

The myth of social factors being more important than physical factors and pay. Central in the work of Mayo is the thesis that social factors are more important than physical factors and pay. Accordingly, Roethlisberger & Dickson (6, p 575–6) concluded that “none of the results [in both relay studies and the mica test] gave the slightest substantiation to the theory that the worker is primarily motivated by economic interest”. The real reason behind the “general improvement” was a change in attitude and morale due to changed supervision (p 190).

Whitehead (5, p 128) also concluded that supervision was this general factor.

This conclusion is in contrast to the data from these studies. After a preferred wage incentive system was introduced, worker output rose (second relay study). After the preferred incentive system was taken away, output promptly dropped (second relay study). When experimental changes were introduced without a change in the regular incentive system, no clear development in weekly output per worker resulted (mica test). When the pay system remained unchanged, performance remained unchanged (bank wiring study). In addition, the interview program underlines the importance of pay factors. In a list of the 15 most prominent company issues “payment” ranked first, whereas “supervision” ranked fourth (table 2).

The “social factors prevail” conclusion has been criticized earlier. Parsons (14, 20) argued that the workers in the first test room experiment were motivated because they systematically received information feedback (ie, knowledge of results about their output rates) and the more individualized piecework pay system. Other scholars performed secondary analyses of the same study’s original data. Franke & Kaul (21) and Franke (22) conducted extensive statistical tests, such as time-series multiple regression analyses. They concluded that, both for the group and for individual workers, the traditional factors “imposition of managerial discipline”, “economic adversity”, and “rest pauses” were responsible for most of the variance in the output scores (21, p 636).

No matter how justified these authors’ appeal for thorough statistical analyses may be, even sophisticated analyses cannot rescue a poor study design. The point is that, even if the methodological shortcomings in the Hawthorne research were waived, neither the original data nor later re-analyses permit the conclusion that social factors are more important than physical factors and pay. On the contrary, especially pay appears to be a key factor in these studies.

The myth of wholehearted cooperation. According to Mayo (8, p 64) the first relay assembly test group “gave itself whole-heartedly and spontaneously to cooperation in the experiment”. As Roethlisberger & Dickson formulated it: “A new supervisor-employee relationship had developed in which there existed a spirit of cooperation with the experimenters and management” (6, p 154). Potential antagonism between management and the Hawthorne employees was denied by Roethlisberger & Dickson: “In the interviews of 1929, where over 40 000 complaints were voiced, there was no single unfavorable comment expressed about the company in general” (6, p 536). The image was created of one happy organizational family and, in the first relay test room, of a committed team that was willing to work ever

harder, irrespective of the physical work conditions and conditions of payment. Nevertheless, clear illustrations of worker resistance and apprehension regarding the experiments were documented by Whitehead (5) and Roethlisberger & Dickson (6). For example, consider the mica test. Only two of the five originally selected operators wanted to join this test group, and operator M1 stated: "The girls all tried to say they were just going to get us in there and time us and then cut the rates, but I thought the other girls who were going were all nice and then, too, it would be quiet in there" (6, p 143–4). Also in the bank wiring study, the men did not trust the investigators. W1 explained: "You know I had the idea that what you people were trying to do was to see if you could get us to do just as much work in six hours as we are doing now". "About everybody down there who has talked about it has the same idea" (6, p 401). These were early reactions, but, also in the course of the first relay test, many indications of antagonism between the operators and management have been documented. Operators 1a and 2a faced disciplinary action (5, p 111–9) and were eventually dismissed from the test group (25 January 1928). From Whitehead (5, p 118), we learn that not only their relation with supervisors was bad, but also their relations within the group. On January 23rd, operator 4 and the lay-out operator requested that either they or operator 2a be removed from the test room. In addition, the supervisors did not always live in harmony. Whitehead (5, p 111–2) reported a struggle between the supervisor of the first test room and his superior about the acceptability of the usual talking of the five operators. A final example: the second relay test ended because the women in the regular department demanded the same rewards as the experimental group. Bramel & Friend (10, 23) have also pointed to the fact "that abundant evidence of worker resistance at Hawthorne was suppressed in influential accounts of the research" (23, p 860). Unfortunately, over time, this observation—that in itself is correct—has been overshadowed by strong criticisms of the presumed ideological nature of their arguments [eg, "Marxist propaganda" (24–27)].

The myth of the neurotic worker. It is interesting to note that, whenever the investigators referred to "worker apprehension", worker resistance or conflicts between workers and management, such phenomena were explained in terms of the mental health status of the employee(s) or by personal circumstances outside the factory (such as family conditions). Worker resistance was thus transformed into an individual, irrational psychological phenomenon, an indicator of poor individual adaptation, without any clear relation to the direct work environment (7, 28).

Referring to the interview program, Roethlisberger & Dickson concluded that "unbalance in the worker"

was expressed as complaints and grievances (6, p 575). Negative attitudes regarding work conditions and style of management were thus explained away by using terms borrowed from psychopathology, such as "obsessive thinking" (6, p 310, p 575), "personal disequilibrium" (4, p 172), "obsessive response" (8, p 66), "emotional blockage" (8, p 72) and "preoccupations" (6, p 184, 292, 311). However, it does not seem improbable that many of the employees' complaints and grievances were influenced by their previous experiences with time and motions studies, with bosses, engineers, and rate-setters (29). In addition, the "Bolshevik" behavior of operator 2a was explained by her change in mental attitude due to a case of anemia (6, p 170), demonstrating "how an organic unbalance could find expression in criticism of company policy" (6, p 325). In the case of operator 1a, the other dismissed girl, a nonwork explanation is offered: She "had been married; it might be expected that after this event her work had no longer its previous significance, and, being easily influenced, she followed the lead of her friend" (6, p 170). Whitehead admits that this interpretation "overlooks the fact that for some months operator 1a had been the leader in this revolt" (5, p 118).

How could this story emerge and survive?

This paper provides evidence that debunks the Hawthorne studies and the Hawthorne effect. It was shown that the experiments do not meet basic methodological criteria, and I critically assessed five Hawthorne myths. Although this is the first article that unfolds the Hawthorne story as five connected Hawthorne myths and that debunks each of them, this is not the first article that has critically addressed the Hawthorne research. Already in 1953, Argyle (30) concluded that the Hawthorne researchers provided "no quantitative evidence for the conclusion for which this experiment is famous, that the increase of output was due to a changed relationship with supervision" (p 100). Other critical accounts of the Hawthorne research have followed (1, 3, 10, 12, 14, 16, 19–21, 31–33). Some of the critical authors performed secondary analyses on the original data; others interviewed original operators or supervisors. But what most of these contributions have in common is their criticism of the common interpretation of the data in terms of a Hawthorne effect. Notwithstanding these publications, the appendix demonstrates that the story of the Hawthorne effect is still much alive. [See also the report of Olsen (11).] Obviously, previous studies were not successful in putting an end to biased accounts and interpretations. The question thus becomes how this story could emerge, and how it could survive. This issue has been touched upon by some authors (19, 31, 34), but a satisfactory answer has not yet been given.

Therefore, necessarily not without speculation, I address this question in more detail now. I postulate that the following five interrelated factors contribute(d) to the birth and survival of the Hawthorne myths.

1. *The story is too good to be untrue:* The story told—employees improving their output irrespective of the physical work conditions and pay—is too good to be untrue. Over time, the Hawthorne effect has become an urban legend. It could be argued that once you have got the story, you do not need the data to prove it. Even so, once you have heard the story, you will never forget it, especially since this message is echoed by authoritative authors of scientific handbooks, and since it is delivered to students in the beginnings of their studies (“primacy effect”).

2. *The original researchers have been biased and selective in their reports; later generations have been “lazy”:* The Hawthorne researchers were not always successful in distinguishing facts from fiction. Mayo [in the foreword to Whitehead’s report (5, p viii)] states that “The presentation of facts invariably implies something of selection”. His accounts were indeed selective in terms of the information reported. The original authors omitted, downplayed, and de-emphasized certain features of their cases. Especially Mayo (4, 8), in his role of scientific popularizer (35), has been important in this respect (18, p 27). He polished the story to its too-good-to-be-untrue status. As Gillespie (34, p 178) has argued: “Mayo was a social scientist trying to use the factory research as the empirical basis of his social theories”. His writings show a “strong need to present as convincing a case as possible, without confusing the issue with qualifications” (18, p 27). The more popularized text of Roethlisberger (7) expresses the same tendency. Even so, researchers from later generations often did not bother to check the original data. When asked why psychologists have been much too uncritical in accepting the accounts of the original researchers and scientific popularizers, Parsons, quoted by Rice (33), pointed to a certain indolence. He replied “because they are lazy”.

3. *Social factors do matter.* Social factors (eg, style of leadership, social processes within and between groups, social facilitation, and inhibition) are definitely important in understanding employee well-being and performance. In an earlier report (36), I compared seven theories on the psychosocial work environment and concluded that four of these identify social relations at work as one of the key job features. Partly, the popularity of the Hawthorne story reflects the empirical fact that social factors are important in explaining differences in performance and well-being. The point is that it is not only social factors that matter.

The Hawthorne-based scientific recognition of social factors deserves two further comments. First, it illustrates the, at that time, tremendous gap between social scientists and the shop floor. [See also the report of Zickar (37).] It is cynical that the fact that workers too do have attitudes, feelings, and sentiments is presented as a scientific breakthrough. Second, the ground-breaking conclusion that workers’ behavior is influenced by more factors than just pay and physical work conditions was not new at all. More than 20 years before, Frederick Taylor (38, p 6) described the “soldiering” or underworking phenomenon: “deliberately working slowly, so as to avoid doing a full day’s work”. Taylor explained that one of the reasons for soldiering was the belief among workers that, if they would work harder, higher performance standards and the loss of jobs would follow. He recognized systematic soldiering as a social phenomenon, stemming from “reasoning caused by their relations with other men” (38, p 10).

4. *The story is in accordance with the cognitive world and interests of psychologists.* We believe that, due to their training, many psychologists, have a tendency to view societal phenomena through the micro-level glasses of psychology. They seek the explanation for behavior in “subjective” factors, such as individual features of the person (eg, within-person processes such as perception). Or they seek explanations in between-person processes (eg, group processes). Such psychological explanations are preferred over explanations in terms of more contextual “objective” characteristics of the environment. The Hawthorne explanation (attention, social influences) fits well in this cognitive scheme. In a similar vein, the practical implications of the Human Relations School (ie, that supervisors should learn social skills, and learn to interview and counsel their employees) suited psychologists well, as it was well in line with their psychological expertise. Obviously, psychologists were important and needed because they had the tools to deal with these problems.

5. *The story is in accordance with the cognitive world and interests of management.* Basically management’s task was twofold, to ensure worker productivity and to control the social process in the factories. In the societal context of the first decades of the 20th century—the United States was a capitalist society characterized by major labor unrest and disputes—the accomplishment of both tasks was not easy. And here was this management philosophy, with a scientific foundation, that claimed that management could reach these two goals. [See also the report of Gillespie (34, p 210 and p 238).] Paying attention to their “unbalanced” workers would bring both goals near, a harmonious and productive factory. The message of Mayo (social factors prevail,

social influences are of utmost significance) even meant that there was no need to redesign work conditions; what really mattered were the social (nonmaterial) concerns! If causes were socially rooted, there was no need to improve work conditions or pay.

Historical studies must be understood in their "at that time" scientific and societal context. [See also the report of Gillespie (34).] Therefore, it would neither be fair nor correct to judge the Hawthorne studies by the scientific and methodological standards of today. Much of today's standard expertise (experimental and quasi-experimental designs, multivariate analyses, theories on work motivation, and group performance) was not available to these pioneers.

Accordingly, it is not the aim to simply criticize the Hawthorne researchers. In fact, in retrospect, one should acknowledge the considerable scope and pioneering nature of the Hawthorne research. It is unique that a factory, in collaboration with universities, built up a strong research infrastructure and maintained it over a long period of time. Especially the interview program constituted an impressive study. Unfortunately, only parts of it have been reported. Moreover, the Hawthorne research program deserves to be commended for its "modern" interdisciplinary approach and combination of study designs: quasi-experimental longitudinal study designs in a natural setting, observational studies, and interview studies. Several potential strong points of the experimental studies deserve to be recognized as well. The following multiple sources of quantitative and qualitative data collection were utilized: performance data, observations, interviews, physiological data (eg, blood pressure, heart rate), and medical examinations. In this sense, the Hawthorne research does indeed constitute a hallmark in the history of social sciences.

Nevertheless, the importance of these studies is seriously overstated in the literature. We have shown that the Hawthorne effect is a myth. One might argue that, although there is no proof of a Hawthorne effect in the original studies, it does not mean that there is no such thing as a Hawthorne effect. However, few people will disagree with the fact that human beings, both in everyday life and when part of a scientific investigation, "reflect upon their situation and react to it when they consider this appropriate". Consequently, there is no need to call this a special effect (3, p 366).

Will there be a future for the Hawthorne effect? I am afraid that there is. Earlier criticism has not been sufficient to put an end to this most persistent myth. Nevertheless, through a thorough account of what really happened in the Hawthorne research, through the formulation and debunking of five Hawthorne myths, and through the postulation of five constituting factors, I hope to have contributed to the unraveling and "unmasking" of the Hawthorne research. At the least, in the fu-

ture, in the teaching of new students, we need to differentiate between fact and fiction.

Acknowledgments

The author wishes to thank Ad van Knippenberg and Toon Taris for their comments on a previous version of this manuscript.

References

1. Adair JG. The Hawthorne effect: a reconsideration of the methodological artifact. *J Appl Psychol.* 1984;78:413-32.
2. Ruch F, Zimbardo P. *Psychology and life.* 8th edition. Glenview (IL): Scott Foresman; 1971.
3. Wickstrom MD, Bendix T. The "Hawthorne effect"—what did the original Hawthorne studies actually show? *Scand J Work Environ Health.* 2000;26(4):363-7.
4. Mayo E. *The human problems of an industrial civilization.* New York (NY): MacMillan; 1933.
5. Whitehead TN. *The industrial worker: a statistical study of human relations in a group of manual workers.* 2 volumes. Cambridge (MA): Harvard University Press; 1938.
6. Roethlisberger FJ, Dickson WJ. *Management and the worker.* Cambridge (UK); London: Harvard University Press; 1939.
7. Roethlisberger FJ. *Management and morale.* Cambridge (MA): Harvard University Press; 1941.
8. Mayo E. *The social problems of an industrial civilization.* London: Routledge & Kegan Paul (reprinted in 1975, with a foreword by JH Smith); 1949.
9. French J. Experiments in field settings. In: Festinger L, Katz D, editors. *Research methods in behavioral sciences.* New York (NY): Holt, Rinehart and Wilson; 1953. p 98-135.
10. Bramel D, Friend R. Hawthorne, the myth of the docile worker, and class bias in psychology. *Am Psychol.* 1981;36(8): 867-78.
11. Olson R, Verley J, Santos L, Salas C. What we teach students about the Hawthorne Studies: a review of content within a sample of introductory I-O and OB textbooks. *Ind Org Psychol.* 2004;41(3):23-39.
12. Gale EAM. The Hawthorne studies—a fable for our times? *QJM.* 2004;97:439-49.
13. Snow CE. Research on industrial illumination: a discussion of the relation of illumination intensity to productive efficiency. *The Tech Engineering News.* 1927: 257, 272, 274, 282.
14. Parsons HM. What happened at Hawthorne? *Science.* 1974;183:922-32.
15. Mulherin J. *The sociology of work and organizations: historical context and pattern of development [dissertation].* Berkeley (CA): University of California; 1980.
16. Greenwood RG, Bolton AA, Greenwood BA. Hawthorne a half century later: relay assembly participants remember. *J Manage* 1983;9(2):217-31.
17. Kompier MAJ, Cooper CL, Geurts SAE. A multiple case study approach to work stress prevention in Europe. *Eur J Work Organ Psychol.* 2000;9(3):371-400.
18. Yorks L, Whitsett DA. Hawthorne, Topeka, and the issue of science versus advocacy in organizational behaviour. *Acad Manage Rev.* 1985;120:21-30.
19. Jones SRG. Was there a Hawthorne effect? *Am J Sociol.* 1992;98(3):451-68.
20. Parsons HM. What caused the Hawthorne effect? A scientific

- detective story. *Adm Soc.* 1978;10:259–83.
21. Franke RH, Kaul JD. The Hawthorne experiments: First statistical interpretation. *Am Sociol Rev.* 1978;43:623–43.
 22. Franke RH. The Hawthorne experiments: re-view. *Am Sociol Rev.* 1979;44:861–7.
 23. Bramel D, Friend R. Is industrial psychology none of Marxism's business? *Am Psychol.* 1982;37:860–2.
 24. Locke EA. Critique of Bramel and Friend. *Am Psychol.* 1982;37:858–9.
 25. Stagner R. The importance of historical context. *Am Psychol.* 1982;37:856.
 26. Feldman J. Ideology without data. *Am Psychol.* 1982;37:857–8.
 27. Vogel W. A Judeo-Christian rejoinder to Bramel and Friend's Marxist critique of the capitalist Hawthorne effect. *Am Psychol.* 1982;37:859–60.
 28. Bourke H. Industrial unrest as social pathology: the Australian writings of Elton Mayo. *Hist Stud.* 1982;20:217–33.
 29. Golden CS, Ruttenger HZ. *Dynamics of industrial democracy.* New York (NY): Harper; 1942.
 30. Argyle M. The relay assembly test room in retrospective. *Occup Psychol.* 1953;27:98–103.
 31. Carey A. The Hawthorne studies: a radical criticism. *Am Sociol Rev.* 1967;32:403–16.
 32. Wardwell WI. Critique of a recent professional 'Put-Down' of the Hawthorne research. *Am Sociol Rev.* 1979;44:858–61.
 33. Rice B. The Hawthorne defect: persistence of a flawed theory. *Psychol Today.* 1982;16:70–4.
 34. Gillespie, R. *Manufacturing knowledge. a history of the Hawthorne experiments.* New York (NY): Cambridge University Press; 1991.
 35. Smith JH. Foreword. In: Mayo E. *The social problems of an industrial civilization.* London: Routledge & Kegan Paul; 1949, reprinted in 1975.
 36. Kompier M. Job design and well-being. In: Schabracq M, Winnubst J, Cooper C, editors. *Handbook work and health psychology.* Chichester (UK): Wiley; 2003:429–54.
 37. Zickar MJ. An analysis of industrial-organizational psychology's indifference to labor unions in the United States. *Hum Relat* 2004;57:145–67.
 38. Taylor FW. *The principles of scientific management.* New York (NY): Harper Brothers; 1911.

Received for publication: 27 April 2006

Appendix

Examples of definitions of the Hawthorne effect (1955–2006)

General psychology and research methodology textbooks

"The most famous example of a laboratory effect is the Hawthorne study, in which selected factory workers were so pleased at the experimental attentions of the management that they produced steadily more with each new change, regardless of whether it was an improvement. The highest productivity came with the final change, which took away all breaks, hot lunches, and incentive payment." (1, p 340)

"These influences—receiving special attention and acting within changed social relations- produced effects that far outweighed the effects of the experimental treatments the experimenters had planned to study. The effects on actors of being singled out to receive the special attention of being treated as experimental subjects has become known as the Hawthorne effect." (2, p 225)

"No matter what changes were made—whether there were many or few rest periods, whether the work day was made longer or shorter, etc—the women produced more and more telephone relays. Although it is hard to pinpoint the reason for this change, since a number of variables were confounded, it seems quite likely that it was largely due to the fact that the women knew they were in an experiment, felt the special attention, and wanted to cooperate. They knew that the experimenters expected the changes in the working conditions to affect them, so they did. They kept working harder and harder." (3, p 395)

"To the researchers' great satisfaction, they discovered that making working conditions better consistently increased satisfaction and productivity. As the workroom was brightened up through better lighting, for example, productivity went up. Lighting was further improved, and productivity went up again. To further substantiate their scientific conclusion, the researchers then dimmed the light: Productivity again improved! It became evident then that the wiring room workers were responding more to the attention given to them by the researchers than to improved working conditions." (4, p 236)

Work and Organizational and Social Psychology textbooks

"Each time an experimental variable was introduced—whether it improved work conditions or made them worse—the group would express its collective high spirits and involvement in the experimental task by increasing production." (5, p 36)

"Regardless of the conditions, whether there were more or fewer rest periods, longer or shorter work days, each experimental period produced a higher rate of work than the one before—the women worked harder and more efficiently." (6, p 414)

"No matter what level the lighting was set at, productivity increased! There was a steady increase in workers' output following any change in lighting. In other studies, Mayo systematically varies the length and timing of work breaks. Longer breaks, shorter breaks, more or fewer breaks, all resulted in a steady increase in worker output." (7, p 9)

"The researchers were surprised to find that over the course of the experiment, productivity increased and seemed to have little to do with the lighting levels. Many explanations of these results have been advanced and debated. The most frequently discussed is that knowledge of being in an experiment, or what has come to be called the Hawthorne effect, caused increases in performance." (8, p 12)

References

1. Blalock HM, Blalock AB. *Methodology in social research*. New York (NY): McGraw-Hill; 1968.
2. Runkel PJ, McGrath J. *Research on human behaviour: a systematic guide to method*. New York (NY): Holt, Rinehart and Winston Inc; 1972.
3. Kantowitz B, Roediger HL. *Experimental psychology*. Chicago (IL): Rand McNally; 1978.
4. Babbie E. *The practice of social research*. Belmont (CA): Wadsworth Publishing Company; 1995.
5. Katz E, Lazarsfeld P. *Personal influence: the part played by people in the flow of mass communication*. New York (NY): Free Press of Glencoe; 1955.
6. Freedman JL, Sears DO, Carlsmith JM. *Social Psychology*. 3rd ed. Englewood Cliffs (NJ): Prentice-Hall; 1978.
7. Riggio RE. *Introduction to industrial/organizational psychology* (4th edition). Upper saddle River (NJ): Pearson Education; 2003.
8. Spector PE. *Industrial and organizational psychology, research and practice*. 4th ed. John Wiley and Sons, Inc, 2006.