Is there a Hawthorne effect?

François Geoffroy,
CRIISEA (EA 4286), IAE Amiens

Original article in French published in *Gérer & Comprendre*, March, 2019, pp. 42-52.

The experiments conducted by Elton Mayo and his team from Harvard at Western Electric's Hawthorne factory led to the formulation of the well-known "Hawthorne effect": the staff's attentive attitude staff motivated wage-earners' behaviors. However more and more studies have questioned these findings and conclusions. This critical, historical reinterpretation of these experiments is based on a review of the literature in several disciplines.

Six experiments were conducted from 1924 to 1933 in five different departments at Hawthorne Works (Cicero, Illinois), where Western Electric employed approximately 29,000 people to make telephones and the switches and cables for them. The objective of these six experiments, some of them carried out simultaneously, was to investigate the effects of fatigue, monotony and lighting on worker productivity. Homer Hibarger, a white collar at the factory, and Charles Snow, a researcher from Massachusetts Institute of Technology (MIT) performed the first experiment. The five others were undertaken by a team from Harvard (Mayo, Roethlisberger, Whitehead, Whyte and Warner) starting in April 1928. This progrjam was under the control of Western Electric's staff (Pennock, Hibarger, Wright and Dickson). The reports were not analyzed or published till a few years later by Whitehead (1938) and then in more detail by Roethlisberger and Dickson (1939).[1]

Since then, these experiments have become well known for their finding of the "Hawthorne effect", which refers to the improved productivity of wage-earners owing to the mere factor that the staff shows more interest in their work. For the school of human relations and, more broadly, the theory of organizations, the Hawthorne experiments made history, since they took the human variable into account and refuted scientific management, *i.e.*, Taylorism (HASSARD 2012).

This research, owing to its opacity, has aroused many reactions. Authors have called it a "*myth*" (BRAMEL & FRIEND 1981, BERT 1999), "*flawed theory*" (RICE 1992), "*fable*" (GALE 2004) and even "*urban legend*" (KOMPIER 2006). However the studies discussing these experiments have not always fully presented them. Olson *et al*. (2004) have analyzed what the 21 textbooks on organizational psychology that rang up the highest sales on Amazon.com in 2003 had to say about these experiments. All these textbooks mentioned the experiment on lighting, thirteen described the "*first relay assembly test room*" experiment; twelve, the "*bank wiring observation room*" experiment; and six, the "*interviewing program*"; but none of them mentioned the "*mica splitting test room*" and "*second relay assembly*" experiments.

We are forced to admit that the managerial sciences and (even more) human resource management have remained impervious, out of naivety or denial, to the criticisms directed at these experiments. So, we can legitimately ask: is there a Hawthorne effect and should we still be talking about it?

This article reviews all six experiments and, too, the controversies and criticisms surrounding them. After presenting the experiments chronologically so as to follow their "logical" succession, both their findings and conclusions will be brought under scrutiny. A review of the literature in several disciplines will then serve as the grounds for rejecting the existence of a Hawthorne effect.

A myth is born: The artificial lighting experiments (November 1924 – April 1927)

On the initiative of the National Research Council of the National Academy of Science,[2] and with funding from General Electric, the manufacturer of electric light bulbs, four series of experiments (called "*illumination studies*") were carried out on artificial lighting at the Hawthorne Works between November 1924 and April 1927. At the time, the only lighting in the workshops came from the windows and/or skylights, a source of problems

[1] This article has been translated from French by Noal Mellott (Omaha Beach, France). The translation into English has, with the editor's approval, completed a few bibliographical references.

[2] Frank Jewett, who headed Western Electric's workshops, became president in 1923 of the engineering division of the National Research Council and convinced them to choose Western Electric (DESMAREZ 1986).

during bad weather. The goal was to show that improving the lighting would increase wage-earner productivity. Besides this material condition other factors came into play, namely: the friendliness of management and piece-rate pay.

Conducted by Snow in November 1924, the first experiment lasted five weeks. This "*illumination test I*" was performed in three departments where ceiling lights were installed. Each department had a test (N=9) and a control (N≈30) group; and a "normal" lighting of 5 fc (foot candles, the standard unit of lighting in the United States). Varying the lighting in the test group (from 3 to 46 fc) had no significant impact on productivity; the results were scattered, varying from service to service (SNOW 1927, PENNOCK 1930, WREGE 1976).

A second experiment ("*illumination test II*") was conducted from December 1924 till the following summer. The participants, all of them from the same department, were divided into a test and a control groups, each with ten members. Each group was placed in an experimental workshop with artificial lighting. The intensity of the light varied from 24 to 70 fc in the test group while, in the control group, it remained more or less constant (16-28 fc depending on the season and time of day). There was no group effect; production increased in the same proportion. However there was a problem with this experiment: the lighting could not actually be controlled owing to daylight.

In 1926, Hibarger, with Snow's help, decided to carry out another experiment ("*illumination and psychological test*") with artificial lighting alone so as to avoid this bias. The windows were made opaque. Three groups of ten participants were formed, a control group with constant lighting (5 fc), a test group with lighting that dimmed by 1 fc (from 10 to 3 fc), and a "psychological" group who were led to believe that the lighting varied whereas it remained constant (In the presence of participants, the light bulb was replaced with another bulb but of the same intensity). Referring to the results from this last group, Hibarger argued that the increase in production was to be set down to direct supervision alone. However Snow (1927) reported that production had not increased but, in fact, had fallen in all groups. According to Roethlisberger and Dickson (1939, pp. 14-18), productivity increased very slightly in the three groups. Given these contradictory conclusions, doubts have never settled. Roethlisberger and Dickson did not take part in this experiment, and their report is a secondary, cursory source that devotes but four out of 604 pages to the lighting experiments. Apparently, Snow's conclusions were more accurate, since he was a researcher involved in this experiment. The results for the "psychological group" were ambiguous. In 1956, Hibarger admitted that no subterfuge had been used in this group (WREGE 1976).

From these three experiments, Snow (1927) concluded that making the lighting dimmer or brighter had no effect on productivity. For this reason, he and the members of the National Research Council gradually withdrew from this program in 1926. According to Snow, the explanation of any variation, when it occurred, could be set down to: *1)* supervisors' pressure on employees; *2)* physiological (*e.g.* headaches, tiredness, etc.) and psychological (*e.g.* daydreaming, coming vacations, etc.) factors; and *3)* the family environment.

In 1927, Hibarger decided to informally conduct his own experiment with two women workers from assembly in the test group and two others in the control group. This "*moon-light test*" had eleven phases (*cf.* Table 1).

Table 1:

Moon-light test (on lighting)

Phase	***Changes made***		***Period (1927)***	***Output***
	Test group	***Control group***		
1	Normal conditions in the department		28 January – 3 February	100%
2	Moved to experimental workshops		4 February	Increase
3	Morning: 1.0 fc Afternoon: 0.39 fc	5.0 fc	5 February	Increase
4	0.200 fc	5.0 fc	6 February	Increase
5	Morning: 0.200 fc Afternoon: 0.102 fc	5.0 fc	7 February	Slight decrease
6	0.080 fc	5.0 fc	/	Stable
7	0.060 fc	5.0 fc	/	Sharp drop
8	1.4 fc	5.0 fc	28 February	Slight decrease
9	Switched experimental workshops		1 March	Increase for both groups
	11.0 fc	5.0 fc		
10	The supervisor's office placed in the workshop		21 March – 9 April	119.5% & 117%
	11 fc	11 fc		
11	The supervisor's office still in the workshop		9 – 23 April	119.5% & 117%
	1 week	2 weeks		

Source: Wrege (1976, pp. 14-15).

During the first phase, the women stayed in their department so that researchers could gauge their individual productivity. During the following phase, they moved to an experimental workshop. The first day that lighting was varied (phase 3), it was set at 1.0 fc in the morning and 0.39 fc in the afternoon for the test group, whose productivity fell slightly lower than the control group's. The next day, the lighting was dimmed to 0.200 fc all day long; and the test group's productivity increased. During the fifth phase, the intensity was reduced to 0.102 fc in the afternoon, and production fell off slightly. When it was dimmed to 0.060 fc (the equivalent of moonlight, whence the experiment's name), the test group's productivity drop (phase 7); and the women objected. During the next phase, the lighting was slightly increased. As of the ninth phase, the test and control groups switched rooms; and the lighting for the test group was now 11.0 fc. The productivity of both groups increased. During the tenth phase, Hibarger's office was placed in the middle of the workshop for a few weeks. The rate of production increased, steadying at 119.5% for the test group and 117% for the control group. The last phase confirmed these results. Phases 9-11 demonstrated the importance of direct supervision on productivity.

What to conclude from the tests on lighting? Throughout this series of experiments, it is clear that the intensity of electric lighting was not related to output. But in reference to the results of the last phases, Hibarger suspected that supervision did have a positive impact. Many questions were left hanging. Why did production decrease during the afternoon? Was it lower Monday and Saturday? Was fatigue a factor? Should breaks be introduced to relieve fatigue? Should the workweek be shortened? To answer these questions, Hibarger set up a new experiment, the "*first relay assembly test room*", that would become a flagship for the school of human relations.

The first relay assembly experiment (April 1927 – June 1932)

The work of assembling relays was tedious, a task taking from 40 to 50 seconds. The women doing this had to assemble 35 parts in a relay switch box. Each worker managed to do about 50 boxes per hour, and had an average output of 2400 per week.

Hibarger formed a group of six workers by selecting two women who were known for their ability to maintain friendly relations and then asking them to choose four others by affinity. Five of the six assembled the relays while the sixth supplied them with the components. After having assembled a relay, it was placed in a hole located at the level of the worker's right hand. It went down a slide that activated a system for registering it and the time between each relay. This system, which made a click for each relay registered, was used to set wages. The group was separated in the test room from other wage-earners on assembly for 270 weeks, a period divided into 24 phases, a change being made during each phase (*cf*. Table 2). Before each change, the women were asked to come to the supervisor's (Pennock's) office where they were told about the new conditions and asked to continue working as usual without paying attention to the change. Observers were constantly present to gather information, both objective (about timetables, breaks, types of relays, the number of relays produced, the quantity of defective items…) and subjective (*e.g*., participants' and observers' comments) (*cf*. Table 2).

During the first phase, the women stayed in their department so that researchers could gauge their individual output. They were then placed in a work group and moved to the test room. Official policy was to tolerate conversations in the test room. Starting with the third phase, the pay system used in the rest of the

Table 2:
The first relay assembly experiment

Phase	*Changes made*	*Period (1927-1929)*	*Production (relays per week)*
1	Normal conditions as in the department	25 April – 10 May 1927	2400
2	The test room + work in a group	10 May – 11 June	2400
3	Financial Incentives based on the group's work	13 June – 6 August	2500
4	Two 5-minute breaks	8 August – 10 September	Increase
5	Two 10-minute breaks	12 September – 8 October	Sharp rise
6	Six 5-minute breaks	10 October – 5 November	2400
7	Two breaks (15 and 10 minutes) + snacks	7 November 1927 – 21 January 1928	2500
8	Same conditions as in phase 7, but the workday ended at 4:30 p.m. (instead of 5:00 p.m.)	23 January – 10 March	Sharp rise
9	Same conditions as in phase 7, but the workday ended at 4:00 p.m.	12 March – 7 April	2900
10	Same conditions as in phase 7, but the workday ended at 5:00 p.m.	9 April – 30 June	2800
11	Same conditions as in phase 7, but without work on Saturday morning	2 July – 1 September	Stable
12	Return to the conditions of phase 3: all benefits eliminated	3 September – 24 November	2900
13	Same conditions as in phase 7 but without snacks	24 November 1928 – 29 June 1929	3000

Roethlisberger and Dickson (1939) reported only thirteen phases for this experiment, since the findings in later phases were of such poor quality that they were not methodically recorded.

plant was modified for the women in the test room. Their wages were measured by the group's average output (and not by the average output of the approximately hundred workers in the department), the goal being to encourage participants to cooperate fully and seriously in the experiment. During phases 4-7, short breaks were introduced when output decreased (end of the morning and start of the afternoon). This decrease, it was observed, was related to hunger, since the women skipped a meal. For this reason, the company offered snacks (fruit, sandwiches, soup, etc.) during the breaks in phases 7-12. Once the breaks were introduced, the test room's results were so convincing that the staff allowed breaks in the whole assembly department as of February 1928. During phases 8, 9 and 11, worktime was shortened. During the tenth phase, Western Electric's staff, surprised by the results, called in two academics: Clair Turner, an ergonomist from MIT, and Elton Mayo, a psychosociologist from Harvard. During the eleventh phase, Saturday morning work was suspended following an agreement with the participants. For them not to be financially penalized, the wages for that morning would still be paid. Despite this good news, production did not budge. During the twelfth phase, researchers took away the previously granted benefits (breaks, snacks, shorter hours). Nonetheless, productivity increased, and this increase lasted. In effect, informal arrangements were soon made during this phase to maintain the group's output: if one woman worked more slowly, another (usually her neighbor) would increase her pace of work to maintain benefits.

The staff gathered the women's opinions in order to know what had motivated them the most:[3] working in a small group; management's friendliness; the pay system; the novelty of the situation; the participant's interest in the experiment; or the attention that the staff and researchers paid to the workers. Turner (1933), while arguing in favor of the impact of financial incentives on the group's productivity, considered that it was but one among other factors. In contrast, Roethlisberger and Dickson (1939) concluded that the staff's friendliness toward the workers fully explained these positive findings.

From the fourteenth phase onwards, the quality of findings deteriorated, and the results were no longer analyzed. The workers cooperated less and less with each other. In 1932, the experiment came to an end, leaving many gray zones behind.

Critique: Critics of this experiment have claimed that the increase in productivity was questionable, that management was authoritarian and that the dedication of these women workers was far from disinterested. Let us examine these three points.

The discussion about whether productivity had increased and about the size of any increase has been heated. Some authors have statistically shown that production did not rise and that there was, therefore, no Hawthorne effect (FRANKE & KAUL 1978, PITCHER 1981, JOHN 1992). In a similar vein, Kompier (2006) has pointed out that, during the twelfth phase, the hourly production rates of four of the five workers definitely fell. This means that average hourly production and total weekly production had been confused. He has also pointed out another anomaly: Roethlisberger and Dickson (1939) and Roethlisberger (1941) do not present the same data.

For others, the increase in output was very small and could be set down to learning (PITCHER 1981, BERT 1999) or standardization of the relays (LECUYER 1988 & 1994, GILLESPIE 1991). The women working in the test room only assembled five sorts of relays, unlike in the assembly department (where 150 different models meant putting together from 26 to 52 parts). The increased output might have followed from the adoption of a "strategy" by wage-earners, since they had been informed about each change prior to its introduction and about the expected results. The women were also able to control the regularity of their output through the sound made by the recording machine (PARSONS 1974, CLAUS 2007). Gillespie (1991) has argued that the workers chose the most favorable changes (pauses and snacks) to increase their output, a conclusion shared by Lecuyer (1994), who has observed that the women had set up a rotation system in which each of them, in turn, was to achieve the daily quota of relays.

Interpretations of interactions in the test room have raised doubts about the sympathetic atmosphere there and management's friendliness (CAREY 1967, BRAMEL & FRIEND 1981). During the third phase, four of the six workers were chided for their lack of "*sincere cooperation*" and their chatting, even though chatting was officially tolerated. During such distractions, productivity fell off significantly. This led to restoring an authoritarian management that forbade chatting; and two of the workers were ousted from the experiment. Mayo (1945) mentioned one dismissal, while the other woman left the test owing to her anemia and not because she had chatted too much. Another reason that might have come into play: the two had the lowest productivity, and management wanted to replace them (CAREY 1967, WALTER-BUSCH 1990, SARIN 2003). Their replacements were two "zealous" workers, one of whom was the fastest on assembly. She was her family's breadwinner and urged the group day after day to make more relays. She assumed the authoritarian role of leader; and her agreement had to be obtained for absences, since the replacement's lack of experience would automatically lower output and, therefore, the bonus. Thanks to the increase in production under this worker's leadership, management could become friendly again (CAREY 1967). In other words, positive results make management friendly — not the reverse, as is often said.

Did financial incentives significantly stimulate productivity? Two workers (Theresa Layman and Wanda Blazejack) and an observer (Donald Chipman) in the Hawthorne experiment were interviewed years later (Parsons 1974 & Greenwood *et al*. 1983). All three stated that financial incentives accounted for the

[3] Miss Rousseau, on staff at the Hawthorne plant, conducted 124 nondirective interviews in 1931 and 1932 (WALTER-BUSCH 1990).

increase in production, since the workers could thus double their pay.

Had there been a Hawthorne effect, then the work group's productivity should have increased with each change in the conditions made during the experiment. Upon scrutiny of the results however, productivity increased very little or even decreased when changes were introduced, even when taking into account a hypothetical lag between cause and effect. In contrast, this increase, when it did occur, can be set down to Taylorist factors such as the standardization of relays, managerial discipline and employees' motivation to earn more pay.

Minimizing the importance of financial incentives compared with social factors became a leitmotif both for the staff and for the Harvard research team. For this reason, a "*second relay assembly test room*" experiment was undertaken.

The second relay assembly experiment (August 1928 – March 1929)

Five new workers were appointed by management for this experiment; they stayed in the assembly department; and working conditions were the same as in the rest of the factory (*cf*. Table 3).

No change was made during the first phase, a period for gauging the average productivity of the five women. During the second phase, researchers explained to them that the pay system would now be based on their group performance (instead of the performance of the whole department). Production immediately jumped 12.6%. This system aroused the jealousy of other workers in the department, who wanted the same pay. During the third phase, under pressure from Pennock, researchers reintroduced the former pay system in order to soothe tensions; and the output from the five workers fell 16.4%. Western Electric then stopped the experiment.

Critique: What to conclude from this aborted experiment? Even though it demonstrated the effect of pay on productivity, Roethlisberger and Dickson (1939) considered that this effect was limited due to the experiment's brevity and the fluctuating productivity of each worker (*e.g*. R3: 117.4% and then 64.7%) and from one worker to the next. For them, two factors explained the increase in production:

- Competition between the first and second relay assembly experiments. The women wanted to do as well as those in the first experiment. This interpretation has no serious grounds. It was based on an informal discussion with the woman who supplied the parts in the first experiment but who did not have a full part in the second. Furthermore, this effect due to competition was not observed during the third phase, when the employees reduced their output to express their discontent.
- The fear of dismissal during hard times in the economy. Doubting this explanation, Gillespie (1991) thought that the 1929 crash had no impact on the experiment. Aware of the importance of this experiment for management, the women knew that they could not be laid off.

Despite Roethlisberger and Dickson's explanations, we are forced to admit that wages were the overriding factor in productivity. As proof, Carey (1967) noted that financial incentives produced, in five weeks, the same increase in production that took nine months in the first relay assembly experiment.

Unable to complete their experiment, researchers designed a new one in a calmer setting: the "*mica-splitting test room*".

The mica-splitting experiment (August 1928 – September 1930)

In the mica-splitting experiment, conditions were the same as in the first (breaks, snacks, shorter hours, etc.) and second (pay incentives) relay assembly experiments. Researchers added piece rate wages, overtime and work on Sunday with higher wages. Their goal was to show that the breaks and a reduction of worktime would but moderately affect production.

Management selected two women workers who then choose three others by affinity. This group of five was then moved to the test room. Splitting and calibrating mica chips was a meticulous job. The research, which lasted 107 weeks, had five phases (*cf*. Table 4). The phases were deliberately longer than in the preceding experiments, the aim being to measure the impact of changes over time. Each woman's production and opinions were recorded (*cf*. Table 4).

The first phase gauged the output for each worker while fostering cooperation between the five. A meeting was organized prior to the movement to the test room in order to explain the objectives of the research and to ask the women to work normally. On this occasion, they were offered to put in overtime. All were in favor of this change, introduced during the second phase. Chatting was tolerated, and the atmosphere was friendly. During the third phase, breaks and work on Sunday were introduced, and immediately accepted

Table 3:

The second relay assembly experiment

Phase	*Characteristics*	*Period (1929)*	*Relays produced*
1	Normal conditions	27 August – 29 September	1634 (100%)
2	Pay based on the group's output	26 November – 26 January	1840 (112.6%)
3	The former pay system	27 January – 14 March	1366 (83.6%)

Table 4: The mica-splitting experiment			
Phase	***Changes made***	***Period (1928-1930)***	***Output %***
1	Normal conditions in the department	27 August – 20 October 1928	100%
2	Move to the test room: overtime (no breaks)	22 October – 24 November	115.6%
3	Overtime; two 10-minute breaks with snacks; Sunday work	26 November 1928 – 15 June 1929	115.6%
4	No overtime; two 10-minute breaks with snacks; Sunday work	17 June 1929 – 17 May 1930	104.4%
5	8 hours per day, 5 days a week; two 10-minute breaks with snacks; but no Sunday work	19 May – 13 September	104.4%

by all. After a few months however, the women took a strong disliking to Sunday work despite the higher wages. During the fourth phase, overtime was done away with, since part of the mica-splitting had been transferred to a Western Electric factory in New Jersey. In August 1929, the rumor was circulating that the whole mica department would be moved to the New Jersey plant. On 16 September 1929, the women in the mica department were all reassigned to other services. However the test room experiment was spared, but overtime and Sunday work were eliminated. This reduction in worktime was accepted. The women, with resignation, suspected that the experiment would soon come to an end.

Critique: What to make of this experiment? For two years, output did increase but variably according to the figures cited by the authors: 120% for all phases according to Pennock (1930) but 115.6% during the second and third phases for Roethlisberger and Dickson (1939) before falling to 104.4% at the end of the experiment. According to Roethlisberger and Dickson, the increase in production was to be set down not to the workers' financial motivations but to the introduction of breaks, while the decrease as of the fourth phase could be explained by the fear of dismissals following the rumor about the mica department.

Although Roethlisberger and Dickson (1939) initially wanted to minimize financial incentives, they forgot that, according to their own explanation, productivity started falling during the fourth phase — when overtime was eliminated, a change that led to a loss of wages. This demonstrated, once again, the significance of pay in the commitment of workers to doing their jobs.

Pennock and the staff had very pragmatic objectives. They wanted to show that the experiment was a success and that the breaks had a positive impact. Western Electric's management was concentrating on factors that it could manage at little cost. For this reason, it fully approved the breaks since, unlike financial incentives, introducing breaks entailed scant, if any, extra costs.

Individual interviews with workers (September 1928 – early 1931)

To complete the preceding experiments and understand the impact of a sympathetic management on productivity, an "*interviewing program*" was set up in the factory under the joint direction of Whitehead, a researcher, and Wright, a staff member. The interviews did not initially yield much information. The questions were precise (*e.g.* How does your supervisor treat you? How do you feel about your working conditions?); and the interviewers (foremen or staff members) were not trained for this assignment.

In July 1929, a new series of less directive interviews, each lasting one and a half hours on the average, was begun that allowed interviewees to freely voice their opinions (with the guarantee of anonymity). In all, a wide range of information, including approximately 40,000 grievances, was garnered from 21,126 interviews. A content analysis identified 74 themes, which were then condensed into 37 categories. Table 5 presents the principal findings.

Table 5: A content analysis of the interviews		
Rank	*Theme*	*Frequency*
1	Pay	6816
2	Cloakrooms	3540
3	Health and safety	3208
4	Direct supervision	2737
5	Overtime	2273
6	Toilets	2044
7	Lighting	1689
8	Ventilation	1524
9	Output quotas (*bogey*)	1384
10	Employment	1318
Source: Roethlisberger & Dickson (1939, p. 232)		

Critique: According to the findings, pay was the most frequently mentioned theme in the interviews. However management attached little importance to criticisms, since it felt that the individual interviews merely echoed personal grievances. Why was it so obstinate? Why did management ignore this theme, so obviously important given its rank? The reason had to do with its proclaimed goal: show that a paternalistic, sympathetic management was the single source of motivation for the personnel — in other words, provide proof of the Hawthorne effect.

To gauge the group dimension, six workers were subjected to direct observation in May 1931. Two research assistants interviewed each one and then observed their activities and group interactions in the workshop for several weeks. However researchers realized that worker behavior could not be understood if the informal organization were overlooked that imposes norms and regulates group activities. For this reason, this observation period came to an end in 1931 in order to make room for a broader program: the "*banking wiring observation room*" experiment.

The bank wiring room experiments (June 1931 – February 1933)

Mayo (1945) assigned this experiment on the assembly of telephone switching equipment to an ethnologist from Harvard, Lloyd Warner. A new test room was installed for observing social interactions within a group of fourteen skilled workers (June 1931 – May 1932): nine assemblers (W1-W9), three welders (S1, S2 and S3 who was replaced with S4), two quality controllers (I1 and I2, the latter replaced after two weeks with I3). These workers, chosen for their strength, were observed for more than six months by someone who remained as "invisible" as possible and had instructions to never intervene and to take note of any words or deeds related to social relations within the group and its informal organization. In parallel, an interviewer conducted interviews for knowing the participants better (*e.g*. their family situation, social life, etc.). The room was organized such that working conditions, in particular the pay system, were identical to those in the department.

The workers had to reach the "bogey", a quota set by management. Above this quota, they were paid a piece rate. Since the supervisor was not constantly in the room, the group enjoyed a degree of autonomy. Two informal subgroups formed: on the one hand, W1, W2, W3, W4, S1 and I1, and on the other hand, W6, W7, W8, W9 and S4. W5, S2 and I3 did not belong to either subgroup. An informal parallel organization spontaneously arose with its leaders, interests, production standards and regulations. No improvement was observed in productivity, even though the workers obviously could have produced many more units (15-20%) without much more effort and would have been paid for the units produced above the quota (ROJOT 2005). The workers seemed to be completely indifferent to financial incentives. According to them, increasing productivity would spur management to increase the bogey.

An informal group process regulated worker production: those who failed to reach the quota were said to be profiteering (the assemblers W1, W8 and W9) and those who overshot the quota were called bootlickers (W2 and W6). This social climate usually forced the profiteer or bootlicker to leave the experiment. Under an informal code of honor, members were never to discredit another member or to "snitch" on the group to management, lest they be physically punished. The two subgroups respected each other but also regularly came into conflict for absurd reasons (such as closing a window opened by a member of the other subgroup). However they stuck together opposite nonmembers, such as supervisors, foremen or the person who supplied parts (whom they called Goofy).

The experiment was officially stopped on 19 May 1932 because of the slack work done by the participants, but it was unofficially pursued till 8 February 1933. Five of the fourteen workers were fired and replaced with others whose output was much lower than their predecessors'; but the researchers did not take their results into account when analyzing the data or drawing conclusions.

Critique: This experiment raises problems as to its scientific validity. The results, since they deteriorated over time, were not recorded by researchers. Once again, we have reasons for doubting the Hawthorne effect. The facts that have just been presented show that, for lack of any legal or formal authority within a group, informal subgroups form in its stead. The major lesson to draw from this experiment is the impossibility of understanding worker behaviors without taking under consideration the group's informal organization, which imposed restrictions on output and resisted changes.

Conclusion

This article has presented the research that, conducted at Western Electric's Hawthorne Works over a period of nearly ten years, claimed to have discovered the well-known Hawthorne effect, *i.e.*, improving human relations in a firm leads to better economic performance. In fact, the phrase "Hawthorne effect" has been wrongly attributed to Elton Mayo, who never used it. It was forged by Roesthlisberger alone (SARIN 2003, KOMPIER 2006). What role did Mayo actually play? This latecomer to the experiments (at the time of the individual interviews) limited his role to commenting the reports about the experiments and making them known. According to Lecuyer (1994, p. 105), Mayo "*turned down Pennock's offer for him to direct the program of interviews with all workers in the factory, but he would suggest to be paid fees as a consultant and to have his expenses reimbursed*". Bruce and Nyland (2011), more critical, thought that Mayo merely lent his name as a faculty member of Harvard University to legitimate the experiments.

In effect, the experiments conducted at Hawthorne came out of a partnership between Western Electric and Harvard, two highly credible institutions in American society. Harvard wanted firms to open their doors to its researchers so as to eventually amplify its reputation, while the factory officially wanted to provide a human

image of paternalistic management.[4] However the unofficial reasons were more politically motivated: curtail protest movements (After all, four million American workers went out on strike against their employers in 1919), weaken workers' interest in alternatives such as socialism and Marxism, and keep the personnel from joining a union (O'CONNOR 1999). With regard to the last point, Western Electric devoted, between 1933 and 1936, $25,825 to spying on unionized workers (HASSARD 2012).

Several questions have been left hanging that force us to have doubts about any Hawthorne effect. The researchers from Harvard did not take part in the Western Electric experiments before April 1928, a year after the start of the first relay assembly experiment. Besides, the final report by Roethlisberger and Dickson (1939), which contained the results, was written ten years after the experiments. Furthermore, given the limited size of the samples, it is impossible to statistically process the results or to generalize the conclusions, even more so since the results that did not meet up to expectations were omitted from the analysis. Although there was an increase in productivity, it was moderate and did not, in any case, come from improved human relations. It could be fully set down to managerial discipline and financial incentives.

Nevertheless, something did happen at Hawthorne. What effects do the experiments performed at Western Electric actually bring to light? Though undertaken to undermine the foundations of scientific management, these experiments ended up proving the effectiveness of this Taylorism: piece rate wages, discipline, standardization. In fact, several decades earlier, Taylor had already conducted studies that drew attention to breaks and shorter hours as ways to reduce fatigue (FRIEDMANN 1946, BRUCE & NYLAND 2011).

The Hawthorne effect is, therefore, quite clearly a myth that, constructed *ex post*, has lasted and still survives in textbooks, owing to the negligence of scholars who quote the experiments at Western Electric without having consulted the original data (PARSONS 1974, RICE 1992).

[4] An anecdote, reported by Hassard (2012), clearly evinces the hypocrisy of Western Electric's management, who systematically cut off heating in the toilets during the winter, and ventilation during the summer, so that workers would return more quickly to their stations!

References

BERT C. (1999) "L'effet Hawthorne: un mythe des sciences humaines" in P. CABIN (editor), *Les Organisations. État des savoirs* (Auxerre: Sciences Humaines Éditions), pp. 59-65.

BRAMEL D. & FRIEND R. (1981) "Hawthorne, the myth of the docile worker, and class bias in psychology", *American Psychologist*, 36, pp. 867-878.

BRUCE K. & NYLAND C. (2011) "Elton Mayo and the deification of human relations", *Organization studies*, 32(3), pp. 383-405.

CAREY A. (1967) "The Hawthorne studies: A radical criticism", *American Sociological Review*, 32(3), pp. 403-416.

CLAUS C.K. (2007) "B.F. Skinner and T.N. Whitehead: A brief encounter, research similarities, Hawthorne revisited. What next?", *Behavior Analyst*, 30(1), pp. 79-86.

DESMAREZ P. (1986) *La Sociologie industrielle aux États-Unis* (Paris: Armand Colin).

FRANKE R.H. & KAUL J.D. (1978) "The Hawthorne experiments: First statistical interpretation", *American sociological review*, 43(5), pp. 623-643.

FRIEDMANN G. (1946) *Problèmes humains du machinisme industriel* (Paris: Gallimard).

GALE E.A. (2004) "The Hawthorne studies: A fable for our times?", *Quarterly Journal of Medicine*, 97, pp. 439-449.

GILLESPIE R. (1991) *Manufacturing Knowledge: A History of the Hawthorne Experiments* (Cambridge: Cambridge University Press).

GREENWOOD R.G., BOLTON A.A. & GREENWOOD R.A. (1983) "Hawthorne a half century later: Relay assembly participants remember", *Journal of management*, 9(2), pp. 217-231.

HASSARD J.S. (2012) "Rethinking the Hawthorne studies: The Western Electric research in its social, political and historical context", *Human Relations*, 65(11), pp. 1431-1461.

JOHN S. (1992) "Was there a Hawthorne effect?", *American Journal of Sociology*, 98(3), pp. 451-468.

KOMPIER M.A. (2006) "The Hawthorne effect is a myth but what keeps the story going?", *Scandinavian Journal of Work, Environment and Health*, 32(5), pp. 402-412.

LECUYER B.P. (1988) "Rationalité et idéologie dans les sciences de l'homme. Le cas des expériences Hawthorne (1924-1933) et de leur réexamen historique", *Revue de synthèse*, 3-4, pp. 401-427.

LECUYER B.P. (1994) "Deux relectures des expériences Hawthorne. Problèmes d'histoire et d'épistémologie" in J.P. BUILLOUD & B.P. LECUYER (editors), *L'Invention de la gestion. Histoire et pratiques* (Paris: L'Harmattan) pp. 93-117.

MAYO E. (1945) *The Social Problems of Industrial Civilization* (Boston, MA: Harvard University Press).

O'CONNOR H. (1999) "The politics of management thought: A case study of the Harvard Business School and the Human Relations School", *Academy of Management Review*, 24(1), pp. 117-131.

OLSON R., VERLEY J., SANTOS L. & SALAS C. (2004) "What we teach students about the Hawthorne studies: A review of content within a sample of introductory IO and OB textbooks", *Industrial-Organizational Psychologist*, 41(3), pp. 23-39.

PARSONS H.M. (1974) "What happened at Hawthorne?", *Science*, 183(4128), pp. 922-932.

PENNOCK G.A. (1930) "Industrial research at Hawthorne: An experimental investigation of rest periods, working conditions and other influences", *Personnel Journal*, 8(5), pp. 296-313.

ROETHLISBERGER F.J. & DICKSON W.J. (1939) *Management and the Worker: An Account of a Research Program Conducted by the Western Hawthorne Works* (Cambridge, MA: Harvard University Press).

ROETHLISBERGER F.J. (1941) *Management and morale* (Cambridge, MA: Harvard University Press).

PITCHER B.L. (1981) "The Hawthorne experiments: Statistical evidence for a learning hypothesis", *Social Forces*, 60(1), pp. 133-149.

RICE B. (1992) “The Hawthorne defect: Persistence of a flawed theory”, *Psychology Today*, 16(2), pp. 70-74.

ROJOT J. (2005) *Théorie des organisations* (Paris: Éditions Eska).

SARIN E. (2003) *Introduction conceptuelle à la science des organisations* (Paris: L’Harmattan).

SNOW C.E. (1927) “Research on industrial illumination: A discussion of the relation of illumination intensity to productive efficiency”, *Tech Engineering News*, November, pp. 257, 272, 274 & 282.

TURNER C.E. (1933) “Test room studies in employee effectiveness”, *American Journal of Public Health*, pp. 577-584.

WALTER-BUSCH E. (1990) “The sudden restlessness of relay assembly operator 2a”, *Working paper*, *Forschungsinstituts fur Arbeit und Arbeitsrecht*, St. Gallen.

WHITEHEAD T.N. (1938) *The Industrial Worker* (Cambridge, MA: Harvard University Press).

WREGE C.D. (1976) “Solving Mayo’s mystery: The first complete account of the origin of the Hawthorne studies — the forgotten contributions of C.E. Snow and H. Hibarger”, *Academy of Management Proceedings*, 1, pp. 12-16.