

The Hawthorne Studies: A Radical Criticism

Author(s): Alex Carey

Source: *American Sociological Review*, Jun., 1967, Vol. 32, No. 3 (Jun., 1967), pp. 403-416

Published by: American Sociological Association

Stable URL: <https://www.jstor.org/stable/2091087>

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/2091087?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



American Sociological Association is collaborating with JSTOR to digitize, preserve and extend access to *American Sociological Review*

JSTOR

investigation altogether. A physicist echoed his biochemist colleague:

It's hard to find the time to work alone. Usually I'm pleased to get the problem solved, especially if it's the kind of problem you can get a research associate to do some work on.

These observations suggest that individual work might decrease after the prize and that collaboration might be substituted in its place as a means of coping with diminishing time for work. Using the percent of single-authored papers to gauge the extent of individual work, we find that, for the laureates as a whole, it remains much the same (38 vs. 40 percent). Increases are confined to men who received the prize after they were fifty and are not significantly greater than would be expected simply as a result of aging.

In summary, the Nobel prize—with its psychological and social consequences for

the men who receive it—is associated with a variety of changes in work patterns. In the period directly after the prize, the productivity of laureates declines more than the productivity of rank-and-file scientists of the same age, but the social demands of the prize are more disruptive for the laureates who experienced comparatively large increments in standing than for those who were already members of the scientific elite. Changes in collaborative relationships also follow the award; prize-winning collaborations terminate sooner, on the average, when only one co-worker has been given the award than when the award is shared by several men. Laureates find new collaborators—as a group, they are not more apt to turn to individual work after the prize. These changes, neither anticipated nor intended by the Nobel Foundation, testify to the multiplicity of consequences of the most prestigious of all awards in science.

THE HAWTHORNE STUDIES: A RADICAL CRITICISM

ALEX CAREY

University of New South Wales

A detailed comparison between the Hawthorne conclusions and the Hawthorne evidence shows these conclusions to be almost wholly unsupported. The evidence reported by the Hawthorne investigators is found to be consistent with the view that material, and especially financial, reward is the principal influence on work morale and behavior. Questions are raised about how it was possible for studies so nearly devoid of scientific merit, and conclusions so little supported by evidence, to gain so influential and respected a place within scientific disciplines and to hold this place for so long.

THERE can be few scientific disciplines or fields of research in which a single set of studies or a single researcher and writer has exercised so great an influence as was exercised for a quarter of a century by Mayo and the Hawthorne studies. Although this influence has declined in the last ten years as a result of the widespread failure of later studies to reveal any reliable relation between the social satisfactions of industrial workers and their work performance, reputable textbooks still refer almost reverentially to the Hawthorne studies as a classic in the history of social science in industry.

One might have expected therefore that the Hawthorne studies would have been subjected to the most searching and skeptical scrutiny; that before the remarkable claims of these studies, especially about the relative unimportance of financial rewards compared with purely social rewards, became so widely influential, the quality of the evidence produced and the validity of the inferences from it would have been meticulously examined and assessed. There have been broad criticisms of Mayo's approach and assumptions, many of them cogent. They include charges of pro-management

bias, clinical bias, and scientific naiveté.¹ But no one has applied systematically and in detail the method of critical doubt to the claim that there is scientific worth in the original reports of the Hawthorne investigators.

BACKGROUND

The Hawthorne studies comprise a long series of investigations into the importance for work behavior and attitudes of a variety of physical, economic, and social variables. The principal investigations were carried out between 1927 and 1932, whereafter economic depression caused their suspension. The component studies may be distinguished as five stages:

- Stage I: The Relay Assembly Test Room Study. (New incentive system and new supervision).
- Stage II: The Second Relay Assembly Group Study. (New incentive system only).
- Stage III: The Mica Splitting Test Room Study. (New supervision only).
- Stage IV: The Interviewing Program
- Stage V: The Bank-Wiring Observation Room Study.

Stages I to III constitute a series of partially controlled studies which were initially intended to explore the effects on work behavior of variations in physical conditions of work, especially variations in rest pauses and in hours of work, but also in payment system, temperature, humidity, etc.

However, after the studies had been in progress for at least twelve months the investigators came to the entirely unanticipated conclusion that social satisfactions arising out of human association in work were more important determinants of work behavior in general and output in particular

than were any of the physical and economic aspects of the work situation to which their attention had originally been limited.² This conclusion came as "the great *éclaircissement* . . . an illumination quite different from what they had expected from the illumination studies."³ It is the central and distinctive finding from which the fame and influence of the Hawthorne studies derive.

This "éclaircissement" about the predominant importance of social satisfactions at work occurred during Stage I of the studies. In consequence, all the later studies are in important ways subordinate to Stage I: "It was the origin from which all the subsequent phases sprang. It was also their main focal point. It gave to these other phases their significance in relation to the whole enquiry."⁴

Stages II and III were "designed to check on" (and were taken to supplement and confirm) the Stage I conclusion "that the observed production increase was a result of a change in the *social situation* . . . (and) not primarily because of wage incentives, reduced fatigue or similar factors."⁵ Stage IV was an interviewing program undertaken to explore worker attitudes. Stage V was a study of informal group organization in the work situation.

The two later studies (IV and V) resulted directly from conclusions based on Stages I–III about the superior influence of social needs. Observations made in both were interpreted in the light of such prior conclusions. Hence it is clear that, as maintained by Urwick, Stage I was the key study, with Stages II and III adding more or less substantial support to it. The present paper will therefore be limited to a consideration of the evidence produced in Stages I–III

¹ For a review of these charges and criticisms see Delbert Miller and William Form, *Industrial Sociology*, New York: Harper, 1951, pp. 74–83. For a defense see Henry A. Landsberger, *Hawthorne Revisited*, New York: Cornell, 1958. Landsberger's defense is restricted to the report of the Hawthorne studies by Fritz J. Roethlisberger and William Dickson, *Management and the Worker*, Cambridge, Harvard Univ. Press, 1939. Even this report, in Landsberger's view, has "done the field of human relations in industry an amount of harm which, in retrospect, appears to be almost irreparable." Landsberger, *op. cit.*, p. 64.

² George A. Penneck, "Industrial Research at Hawthorne," *Personnel Journal*, 8 (February, 1930), pp. 296–313; Mark L. Putman, "Improving Employee Relations," *Personnel Journal*, 8 (February, 1930), pp. 314–325.

³ Fritz J. Roethlisberger, *Management and Morale*, Cambridge: Harvard University Press, 1941, p. 15.

⁴ Lyndall Urwick and Edward Brech, *The Making of Scientific Management*, vol. III, London: Management Publications Trust, 1948, p. 27. See also Roethlisberger and Dickson, *op. cit.*, p. 29.

⁵ Morris S. Viteles, *Motivation and Morale in Industry*, London: Staples, 1954, p. 185.

for the famous Hawthorne conclusions about the superior importance for work behavior of social needs and satisfactions.

THE PREFERRED INCENTIVE SYSTEM AND OUTPUT

Stage I: Relay Assembly Test Room (new incentive and new supervision). In Stage I of the Hawthorne studies, five girls who were employed assembling telephone relays were transferred from the factory floor to a special test room. Here their output of relays was recorded for over two years during which a large number of alterations were made in their working conditions. These alterations included a much less variable assembly task,⁶ shorter hours, rest pauses, freer and more friendly supervision, and a preferred incentive system.⁷ These changes were introduced cumulatively and no control group was established. Nonetheless, it was originally expected that the study would yield information about the influence of one or another physical condition of work.⁸

At the end of two years, the girls' output had increased by about 30 percent.⁹ By this time, the investigators were confident that the physical changes in work conditions had been of little importance, and that the observed increase was due primarily to a change in "mental attitude" of the employees resulting from changed methods of supervision.¹⁰ This change in mental attitude was chiefly characterized by a more relaxed "relationship of confidence and friendliness . . . such . . . that practically no supervision is required."¹¹

However, the standard report of the study recognizes that any of several changes introduced concurrently could, hypothetically, have caused both the observed change in mental outlook and the associated increase in output. The authors of the report list the following as providing possible "hypotheses to explain major changes" in work

behavior:¹² (i) changes in the character and physical context of the work task; (ii) reduction of fatigue and monotony consequent upon introduction of rest pauses and reduced hours of work;¹³ (iii) change in the payment system; and (iv) changes in supervision with consequent social changes in group relations.

The remainder of this paper will critically examine the evidence and arguments from which the investigators reached conclusions favorable to the last of these alternative hypotheses.

First hypothesis: changes in work task and physical context. The investigators allow that "the fact that most of the girls in the test room had to assemble fewer types of relays could not be entirely ignored. Operator 5's performance offered a convincing example. Of all the girls in the room she had had more different types of relays to assemble and of all the girls her output rate had shown the least improvement."¹⁴ Whitehead reports that "later (1930-31) her (Operator 5's) working conditions were in line with the rest of the group and her comparative standing in the group definitely improved."¹⁵

However, it was subsequently found that statistical analysis of the relevant data (i.e., the varying output of five girls who were subjected to numerous cumulatively introduced experimental changes) did not show "any *conclusive* evidence in favor of the first hypothesis." On this ground the investigators "concluded that the change from one type of relay to another familiar type did not sufficiently slow up output to explain the increased output of the relay test room assemblers as compared with the assemblers in the regular department."¹⁶ This conclusion leads the investigators to dismiss from further consideration the possibility that

¹² Roethlisberger and Dickson, *op. cit.*, pp. 86-89.

¹³ The investigators list fatigue and monotony as separate hypotheses. For brevity, these have been combined as one hypothesis. The same sort of critical objections are relevant to the arguments and evidence advanced by the investigators with respect to both.

¹⁴ *Ibid.*, p. 87.

¹⁵ T. North Whitehead, *The Industrial Worker*, London: Oxford Univ. Press, 1938, Vol. I, p. 65.

¹⁶ Roethlisberger and Dickson, *op. cit.*, p. 89. (Italics added.)

⁶ Roethlisberger and Dickson, *op. cit.*, pp. 21, 26.

⁷ *Ibid.*, pp. 22, 30-73.

⁸ *Ibid.*, p. 129; Pennock, *op. cit.*, p. 299.

⁹ Roethlisberger and Dickson, *op. cit.*, p. 160.

¹⁰ *Ibid.*, pp. 189-190; Pennock, *op. cit.*, pp. 297-309.

¹¹ Pennock, *op. cit.*, p. 309.

changes in task and conditions played any part at all in the observed increase in output.¹⁷

Second hypothesis: reduced fatigue due to rest pauses and shorter hours. The investigators recognize that "the rest pauses and shorter hours (may have) provided a relief from cumulative fatigue" resulting in higher output. They acknowledge that the fact that the rate of output of all but the slowest worker declined once the girls were returned to standard hours is "rather convincing evidence in favor of this argument."¹⁸ Yet the investigators eventually dismiss these factors on the grounds that under the new conditions of work neither work curves nor medical examinations provided evidence that fatigue effects were present. Viteles has commented bluntly in this connection: "It is interesting to note that (these grounds) are exactly the same used by other investigators in illustrating the effectiveness of rest pauses *by reason of reduced fatigue*."¹⁹

By these arguments, the investigators eliminated the first two of the four hypotheses originally proposed as alternative explanations of the 30 percent increase in output observed in Stage I. This left two contending "explanations," the new incentive system, and the new kind of supervision and related social factors. The problem of choosing between these explanations led directly to the next two major experiments.

Stage II: Second Relay Assembly Group (new incentive system only). "The aim of (this experiment) was to reproduce the test-room situation (i.e., Stage I) only in respect to the one factor of method of payment, using another group of operators. Since

method of payment was to be the only alteration from the usual situation, it was thought that any marked changes in output could be reasonably related to this factor."²⁰

Five girls who were employed on the same sort of task as the girls in Stage I under normal conditions on the factory floor were given the preferred incentive system which had been used throughout Stage I. Under this system, the earnings of each girl were based on the average output of the five. Under the regular payment system, the earnings of each girl were based on the average output of the whole department (i.e., about 100 girls).

Almost at once the Stage II girls' output increased by 12.6 percent.²¹ But the experiment caused so much discontent among the rest of the girls in the department, who wanted the same payment conditions,²² that it was discontinued after only nine weeks. The output of the five girls promptly dropped by 16 percent.²³

As Viteles comments, "the increase in output during the period when the wage incentive was in effect, followed by a production decrease with the elimination of the wage incentive, represents evidence ordinarily interpreted as indicative of the direct and favorable influence of financial incentives upon output."²⁴ However, the investigators reject this interpretation and, without producing supporting evidence of any substance, conclude firmly²⁵ that the increase was due to inter-group rivalry resulting from the setting up of this second small group.

The change in payment system alone (Stage II) produced as much increase in output in nine weeks (possibly five weeks²⁶) as was produced in about nine months by change in payment system together with a

¹⁷ The scientifically illiterate procedure of dismissing non-preferred explanations on the grounds that (i) the experimenters had found no *conclusive* evidence in favor of them and/or (ii) there was no evidence that any *one* of these explanations, considered by itself, accounted for *all* the effect observed, recurs throughout Roethlisberger and Dickson's report of the Hawthorne studies. This procedure is never applied to preferred hypotheses, which are assumed to be well-founded provided only that the evidence *against* them is less than conclusive. See, e.g., Roethlisberger and Dickson, *op. cit.*, p. 160 and pp. 96, 108, 127.

¹⁸ *Ibid.*, p. 87.

¹⁹ Morris S. Viteles, *Industrial Psychology*, New York: Norton, 1932, p. 476. Italics in original.

²⁰ Roethlisberger and Dickson, *op. cit.*, p. 129.

²¹ *Ibid.*, pp. 131-132, 577; Pennock, *op. cit.*, p. 307.

²² *Ibid.*, p. 133.

²³ According to an earlier report (Pennock, *op. cit.*, p. 307), the increase in output was 13.8 percent, the experiment was discontinued after five weeks, and output then fell by 19-24 percent.

²⁴ Viteles, *Motivation . . .*, *op. cit.*, p. 187.

²⁵ Roethlisberger and Dickson, *op. cit.*, pp. 133-134, 158, 577.

²⁶ Pennock, *op. cit.*, p. 307.

change to genial supervision (Stage I).²⁷ Yet this comparison appears not to have made any impression on the investigators' confidence about the superior importance of social factors.²⁸

Stage III: Mica Splitting Test Room (new supervision but no change in payment system). In *Stage I*, numerous changes had been introduced, resulting in a 30 percent increase in output. In *Stage II*, only one of these changes (the preferred incentive system) was introduced and a rapid 12 percent increase in output resulted. In *Stage III*, "the test-room situation was to be duplicated in all respects except for the change in pay incentive. If . . . output showed a trend similar to that noted in (Stage I), it would suggest that the wage incentive was not the dominant factor in the situation."²⁹ Stage III, then, sought to test the combined effect on output of change to a separate room, change in hours, and the introduction of rest pauses and friendly supervision. Again a selected group of five girls was closely studied and an increase in output was recorded—15.6 percent in fourteen months³⁰ or, if one follows Pennock, 20 percent in twelve months.³¹

A comparison between Stage III and Stage I has little prospect of scientific usefulness since in Stage III (i) the incentive system was different from both the disliked system used at the beginning of Stage I and the preferred system introduced shortly afterwards, (ii) the type of work was quite different from Stage I, and (iii) the experimental changes were quite different.³² However, it is this comparison which has been taken by reporters of the studies³³ and by textbook authors³⁴ to provide the

principal experimental evidence about the relative importance of financial and social motives as influences on output. Assuming with Roethlisberger and Dickson that Stage I and Stage III have some minimum comparability, it is important to examine precisely how the investigators dealt with the evidence from these stages for the purpose of the comparison.

Comparison Between Results in Stages I, II, and III. (i) Stage III produced a claimed 15 percent increase in rate of output over fourteen months. Thereafter the group's average rate of output declined for twelve months before the study was terminated due to the depression and lay-offs. The investigators attribute this decline *entirely* to anxieties induced by the depression,³⁵ ignoring the possibility that the preceding increase might also have been influenced by changing general economic and employment conditions. They do this despite evidence that output among a group of 5,500 Hawthorne workers rose by 7 percent in the two years preceding the experiment.³⁶

(ii) In Stage III, the output rate for each girl shows continuous and marked fluctuations over the whole two years of the study.³⁷ To obtain the percentage increase to be attributed to each girl the investigators chose, for each girl, a "peak" output period within the study period and measured her increase as the difference between this peak and her output rate at the outset of

and a combination of the two gave an increase of 30%. This looks surprisingly like an additive effect, with the social rewards being somewhat more potent in influencing behaviour than the monetary reward." Ross Stagner, *Psychology of Industrial Conflict*, New York: Wiley, 1956, pp. 131–132. See also Milton Blum, *Industrial Psychology and Its Social Foundations*, New York: Harper, 1949, p. 26.

³⁵ Viteles comments on this period of declining output: "Both 'the investigators and the operators were of the opinion that the rates on the new piece parts were not high enough in comparison with the old.' Nevertheless scant consideration is given to the possibility that . . . a reduced appeal to economic motives could readily account in large part for the very severe drop in output observed during this final phase of the *Mica Splitting Room experiment*." Viteles, *Motivation . . .*, *op. cit.*, p. 191.

³⁶ Whitehead, *op. cit.*, vol. II, Chart J-53.

³⁷ Roethlisberger and Dickson, *op. cit.*, p. 147.

²⁷ That is, by the end of Experimental Period 7 in Roethlisberger and Dickson's output chart, *op. cit.*, p. 78.

²⁸ Roethlisberger and Dickson, *op. cit.*, pp. 160, 577.

²⁹ *Ibid.*, p. 129.

³⁰ *Ibid.*, p. 148.

³¹ Pennock, *op. cit.*, p. 307.

³² Roethlisberger and Dickson, *op. cit.*, pp. 156, 159.

³³ *Ibid.*, pp. 146–149, 159–160; Pennock, *op. cit.*, p. 307.

³⁴ For example, "we cannot avoid being impressed by the fact that a wage incentive alone (Stage II) increased production 12%, a change in the social situation raised output 15%, (Stage III)

the study.³⁸ These peaks occur at different dates for different girls. To secure the 15 percent increase that is claimed, the study is, in effect, terminated at different conveniently selected dates for different girls. There is *no one period* over which the group achieved the 15 percent average increase claimed.³⁹

(iii) In Stage I, two measures of the workers' performance are used: total output per week,⁴⁰ and hourly rate of output by weeks.⁴¹ It is not clear from Roethlisberger and Dickson's report of Stage I whether the increase is in *total output* or *rate of output*. It is described only as "increase in output," and "output rose . . . roughly 30%,"⁴² which would ordinarily be taken to mean an increase in *total output*. But the investigators make it clear in passing⁴³ that throughout the studies they used rate of output per hour as "the most common arrangement of output data" by which to "portray the general trend in efficiency of each operator and of the group." Whitehead, who produced a two-volume statistical study of Stage I as companion volumes to Roethlisberger and Dickson's standard report, is very clear on this point: "All output will be expressed in the form of a *rate* . . . as so many relays per hour."⁴⁴

However, Whitehead employs throughout his study the description "*weekly rate of output*" when he means *rate of output per hour by weeks*.⁴⁵ This practice, coupled with his habit of not labelling the ordinates of his charts dealing with changes in output, and added to by Roethlisberger and Dickson's use of phrases such as "increase in output" to mean both *increase in rate of output per hour* and *increase in total output*, has led to widespread misinterpretation of the Hawthorne results, and textbook accounts which are seriously in error.⁴⁶

Several points are of present importance. For Stage I, it is not clear whether the 30 percent increase in output claimed refers to *rate of output* or *total output*. It does not matter which measure is used to calculate percent increase in output in Stage I since the total hours worked per week at the end of the study period is only 4.7 percent less than at the beginning.⁴⁷ Thus, an increase of the order of 30 percent would result from either method of calculation. In Stage III, however, it makes a great deal of difference which method is used, and hourly rate of output is the only measure used. Thus, the 15 percent "increase in output"⁴⁸ claimed for Stage III is an increase in *rate of output per hour worked*, not in *total output*. Indeed, it is only by this measure that any increase *at all* in output can be shown.

If *total output per week* is used to measure performance in Stage III, the 15 percent increase claimed for Stage III reduces to less than zero because although output per hour increased by 15 percent, the weekly hours decreased by 17 percent, from 55½ to 46⅙.⁴⁹

From Evidence to Conclusions. By subtracting the 15 percent increase in Stage III (which is an increase in *rate of output*) from the 30 percent increase in output in Stage I (which is all, or nearly all, an increase in *total output*), the investigators conclude that 15 percent remains as "the maximum amount (of increase in output) to be attributed to the change in wage incentive" introduced in Stage I. The investigators acknowledge the wholly speculative nature of this calculation, yet go on to assert in a summary of events to date that the conclusion "seemed to be warranted from the test room studies so far . . . that

Harmondsworth: Penguin, 1954, pp. 71-72. These authors incorrectly report an almost continuous increase in total weekly output over the first nine months of Stage I. In fact, there was no increase except in the period of eight weeks immediately following the introduction of the preferred incentive system. There was no improvement in weekly output in either the preceding period or the four experimental periods extending over six months which followed it.

⁴⁷ Roethlisberger and Dickson, *op. cit.*, pp. 76-77.

⁴⁸ *Ibid.*, pp. 159-160.

⁴⁹ *Ibid.*, pp. 136-139.

³⁸ *Ibid.*, p. 148.

³⁹ *Ibid.*, pp. 146-148, 159-160.

⁴⁰ *Ibid.*, p. 78.

⁴¹ *Ibid.*, p. 76.

⁴² *Ibid.*, p. 160.

⁴³ *Ibid.*, pp. 55, 77.

⁴⁴ Whitehead, *op. cit.*, vol. I, p. 34.

⁴⁵ *Ibid.*, vol. II, Chart B4.

⁴⁶ For example, Edwin Ghiselli and Clarence Brown, *Personnel and Industrial Psychology*, New York: McGraw Hill, 1948, pp. 435-437; and James A. C. Brown, *Social Psychology of Industry*,

it was impossible to consider (a wage incentive system) as a thing in itself having an independent effect on the individual.”⁵⁰

It is important to appreciate just how invalid are the inferences made. In Stage I, friendly supervision and a change to a preferred incentive system led to an increase in total output of about 30 percent. In Stage III, friendly supervision without a change in payment system led to no increase in total output, but to a less than compensating increase in output per hour over a period during which working hours were reduced from 55½ to 46⅙. This could be interpreted to mean that when working hours exceed about 48 per week such extra working-time may bring little or no increase in total output—a finding which had been well-established many years before.⁵¹ This interpretation would have left the way clear to attribute the 30 percent increase in Stage I entirely to the preferred incentive system. Instead, by the rather special method of analysis and argument that has been outlined, the investigators reached the conclusion that the effect of a wage incentive system is so greatly influenced by social considerations that it is impossible to consider it capable of independent effect.

A similar situation holds with regard to Stage II. As Stage II was planned, the “method of payment was to be the only alteration from the usual situation” with the express intention that “any marked changes in output” could then be “related to this factor.”⁵² There *was* a marked change in output—an immediate 12 percent increase. There *was* an immediate change in behavior—the other girls in the department demanded the same conditions. This would seem to require a conclusion in favor of

the importance of a preferred incentive system, but no such conclusion was reached.

As a first step in the interpretation of the Stage II results, Roethlisberger and Dickson noticed, *post hoc*, that somewhere in the “daily history record” of the Stage I group was a reference to a comment by one member of that group that a “lively interest” was being taken in their output by members of the new Stage II group.⁵³ At this point, the investigators simply note this and hint at significance to come. Twenty-four pages later we are told that “although output had risen an average of 12% in (Stage II) it was *quite apparent* that factors other than the change in wage incentive contributed to that increase . . . *There was some evidence* to indicate that the operators in (Stage II) had seized upon this test as an opportunity to prove to everyone that they could do as well as the (Stage I) operators. They were out to equal the latter’s record. In view of this, even the most liberal estimate would put the increase in output due to the change in payment alone at somewhat less than 12%.” (Italics added). Since no additional evidence had been produced, this judgment lacks any serious foundation.

Much later (p. 577) the matter is returned to and, with no additional evidence, we are given to understand that the increase in output in Stage II was due to certain “social consequences” of the “basic social situation.” This situation is simply asserted to have been one in which “rivalry (with the Stage I group) was brought to a focus” by setting up the Stage II group whose “output rose rapidly” in consequence.

Stage II was “designed to test the effect of a (change in) wage incentive” on output.⁵⁴ The preferred incentive system was introduced and output immediately rose 12 percent. It was withdrawn and output immediately dropped 17 percent. Not encouraging results for anyone who believed that wage incentives were relatively unimportant and incapable of “independent effects.” Yet these awkward results were not only explained away but converted to positive support for

⁵⁰ *Ibid.*, p. 160. Viteles bluntly rejects this inference as invalid, but textbook treatments of the Hawthorne studies generally accept it without demur. Viteles, *Motivation . . .*, *op. cit.*, p. 193.

⁵¹ Horace M. Vernon, *Industrial Fatigue and Efficiency*, London: Dutton, 1921. Ghiselli and Brown have summarized Vernon’s findings as follows: “In a munitions plant, when the working week was reduced from 66 to 48.6 hours (a reduction of 26%) hourly output was increased by 68% and total output for the week by 15%. This instance could be multiplied many times.” Ghiselli and Brown, *op. cit.*, p. 242.

⁵² Roethlisberger and Dickson, *op. cit.*, p. 129.

⁵³ *Ibid.*, p. 134.

⁵⁴ *Ibid.*, p. 576.

just such conclusions, all on the basis of a single hearsay comment by one girl.

The investigators carry the day for the hypothesis that "social factors were the major circumstances limiting output." They conclude that "none of the results (in Stages I, II and III) gave the slightest substantiation to the theory that the worker is primarily motivated by economic interest. The evidence indicated that the efficacy of a wage incentive is so dependent on its relation to other factors that it is impossible to separate it out as a thing in itself having an independent effect."⁵⁵ This conclusion is a striking contrast to the objective results obtained in Stages I, II, and III as these bear on incentive systems: (i) when a preferred wage incentive system was introduced, total weekly output per worker rose (Stage I and Stage II); (ii) when the preferred incentive system was withdrawn, output promptly dropped (Stage II); (iii) when changes in supervision, hours, etc. were introduced but with *no change in incentive system*, no increase in weekly output per worker resulted (Stage III).

Viteles, an unusually perceptive critic of the Hawthorne studies, has commented caustically on Stage III: "This increase in output, representing an average rise of 15% in the first 14 months of the experiment, would ordinarily be accepted as evidence that the introduction of rest pauses and the shortening of the work day can in themselves result in increased output, even in the absence of changes in the way of enhancing the wage incentive."⁵⁶ Yet Viteles misses the important point that there was no overall increase in total weekly output in Stage III—only a less than compensating increase in output per hour when shorter hours were worked. It is clear that he supposes the 15 percent increase to be an increase in total output.⁵⁷ Viteles' patience is great, and his criticism of the Hawthorne studies restrained. But they eventually draw from him a testy general protest about "the more 'subtle'—certainly more subjective—form of analysis and interpretation which has generally characterized interpretation

of the Hawthorne data by the Harvard group."⁵⁸

It remains to consider more closely the complementary Hawthorne claim that it was friendly supervision and social factors which were the principal influences leading to the large rise in output in Stage I.

A CLOSER LOOK AT FRIENDLY SUPERVISION IN ACTION

The *whole* of the Hawthorne claim that friendly supervision and resulting work-group social relations and satisfactions are overwhelmingly important for work behavior rests on whatever evidence can be extracted from Stage I, since that is the only study in the series which exhibits even a surface association between the introduction of such factors and increased output.

Stage I began with five girls specially selected⁵⁹ for being both "thoroughly experienced" and "willing and cooperative,"⁶⁰ so there was reason to expect this group to be more than ordinarily cooperative and competent. Yet from very early in the study "the amount of talking indulged in by all the operators" had constituted a "problem," because it "involved a lack of attention to work and a preference for conversing together for considerable periods of time."⁶¹ The first indication in the report that this might be a serious matter occurs on August 2nd, 1927, twelve weeks after the girls' installation in the test-room, when four of the five operators were brought before the foreman⁶² and reprimanded for talking too much.⁶³ Until November, however, "no attempt had been made to do away with this privilege, although several attempts had been made by the foreman to diminish what

⁵⁵ *Ibid.*, p. 256.

⁵⁹ Note, however, that while the five girls were "all chosen from among those with a considerable experience in the assembly of this kind of relay" . . . "the actual method of selection was quite informal and somewhat obscure; it appears to have been determined by the girls themselves in conjunction with their shop foreman." Whitehead, *op. cit.*, vol. I, p. 14.

⁶⁰ Roethlisberger and Dickson, *op. cit.*, p. 21.

⁶¹ *Ibid.*, p. 53.

⁶² Foremen were on a par with departmental chiefs and four ranks above operatives. *Ibid.*, p. 11.

⁶³ *Ibid.*, p. 38.

⁵⁵ *Ibid.*, pp. 575-576.

⁵⁶ Viteles, *op. cit.*, *Motivation* . . . , p. 190.

⁵⁷ *Ibid.*, p. 5.

seemed to him an excessive amount of talking." But Operators 1A and 2A in particular continued to fail to display "that 'whole-hearted cooperation' desired by the investigators." "Any effort to reprimand them would bring the reply 'We thought you wanted us to work as we feel'";⁶⁴ since that was what the supervisors had told them at the beginning of the study.⁶⁵

By November 17th, 1927, the situation had not improved and disciplinary rules were resorted to. All of the operators were required to call out whenever they made mistakes in assembly, and they were prevented from talking. By December, "the lack of cooperation on the part of some of the operators was seriously alarming a few of the executives concerned." Supervisors were asked to give the girls a "hint" by telling them that they were not doing as well as expected of them and that if they didn't improve they would lose their free lunches.⁶⁶

From now on the girls, but especially 1A and 2A, were "threatened with disciplinary action" and subjected to "continual reprimands." "Almost daily" 2A was "reproved" for her "low output and behavior" (sic).⁶⁷ The investigators decided 1A and 2A did not have "the 'right' mental attitude." 2A was called up before the test-room authorities "and told of her offenses of being moody and inattentive and not cooperative." She was called up again before the superintendent.⁶⁸ Throughout this period output for all five girls remained static or falling.⁶⁹ After eleven weeks of serious but ineffective disciplinary measures and eight months after the beginning of the study, 1A and 2A were dismissed from the test room for "gross insubordination" and declining or static output.⁷⁰ Or, as Whitehead puts it, they "were removed for a lack of cooperation, which

would have otherwise necessitated greatly increased disciplinary measures."⁷¹

1A and 2A were replaced by two girls chosen by the foreman⁷² "who were experienced relay assemblers and desirous of participating in the test." These two girls (designated Operators 1 and 2) were transferred to the test room on January 25th, 1928.⁷³ They *both* immediately produced an output much greater (in total and in rate per hour) than that achieved by *any* of the original five girls on their transfer to the test room and much above the performance *at any time* of the two girls they replaced.⁷⁴

Operators 1 and 2 had been friends in the main shop. Operator 2 was the only Italian in the group; she was young (twenty-one) and her mother died shortly after she joined the test room;⁷⁵ after this "Operator 2 earned the larger part of the family income." "(F)rom now on the history of the test room revolves around the personality of Operator 2."⁷⁶ Operator 2 rapidly (i.e., without any delay during which she might have been affected by the new supervision) adopted and maintained a strong and effective disciplinary role with respect to the

⁷¹ Whitehead, *op. cit.*, vol. I, p. 118. In Mayo's accounts it is first said that these two operators "dropped out" (Elton Mayo, *The Human Problems of an Industrial Civilization*, Boston: Harvard Business School, 1946, p. 56) and later that they "retired." (Elton Mayo, *The Social Problems of an Industrial Civilization*, London: Routledge and Kegan Paul, 1949, p. 62.) It is also interesting to compare the above account of events in the test room and drawn from the standard reports with Mayo's picture of the test room. According to Mayo's account, success was achieved "largely because the experimental room was in charge of an interested and sympathetic chief observer. He understood clearly from the first that any hint of the 'supervisor' in his methods might be fatal to the interests of the inquiry . . . He helped the group to feel that its duty was to set its own conditions of work, he helped the workers to find the 'freedom' of which they so frequently speak . . . At no time in the (whole period of the study) did the girls feel that they were working under pressure." (Mayo, *The Human Problems . . .*, *op. cit.*, pp. 68-69).

⁷² Roethlisberger and Dickson, *op. cit.*, p. 60.

⁷³ *Ibid.*, pp. 55, 56, 60.

⁷⁴ *Ibid.*, Figure 6, p. 76 and Figure 7, p. 78. Compare output curves during the first seven Experimental Periods with output from the second week of Experimental Period 8.

⁷⁵ *Ibid.*, pp. 61-62.

⁷⁶ Whitehead, *op. cit.*, vol. I, p. 120.

⁶⁴ *Ibid.*, p. 53.

⁶⁵ *Ibid.*, p. 21; Whitehead, *op. cit.*, vol. I, p. 26.

⁶⁶ Whitehead, *op. cit.*, vol. I, p. 16.

⁶⁷ *Ibid.*, pp. 116-118.

⁶⁸ Roethlisberger and Dickson, *op. cit.*, p. 55. Superintendents controlled a branch of the works and were seven ranks above operators. *Ibid.*, p. 11.

⁶⁹ *Ibid.*, p. 78. See Experimental Period 7 in Figure 7.

⁷⁰ *Ibid.*, pp. 53-57.

rest of the group,⁷⁷ and led the way in increased output in *every* period from her arrival till the end of the study. In this she was closely followed by the other new girl, Operator 1.⁷⁸

At the time that Operators 1 and 2 were brought into the test room, daily hours of work were shortened by half an hour but it was decided to *pay the operators the day rate for the half hour of working time lost*. A little later, the working day was reduced by a further half hour, and again the girls were paid for the time (one hour per day) they didn't work.⁷⁹ Later still, the girls were given Saturday mornings off and again they were paid for the time not worked.⁸⁰

Summing up experience in the test room up to *exactly* the time when the two operators were dismissed,⁸¹ the investigators claim that "it is clear" that over this period there was "a gradual change in social inter-relations among the operators themselves, which displayed itself in the form of new group loyalties and solidarities . . . (and) . . . a change in the relations between the operators and their supervisors. The test room authorities had taken steps to obtain the girls' cooperation and loyalty and to relieve them of anxieties and apprehensions. From this . . . arose . . . a change in human relations which came to be of great significance in the next stage of the experiment, when it became necessary to seek a new hypothesis to explain certain unexpected

results of the inquiry."⁸² In view of the evidence reviewed here this would seem to be a somewhat sanguine assessment of developments in the test room up to this point. It is, therefore, necessary to examine more systematically the way in which the behavior of the supervisors on the one hand and of the operators on the other (including their changing output) varied during the period under consideration.

It is already clear that whatever part satisfying social relations at work—resulting from free and friendly supervision—may have played in producing the increase in output, there were other influences likely to have been important, e.g., a period of fairly stern discipline, the dismissal of two workers, and their replacement by people of rather special personality and motivation. In order to assess these various influences on output it is necessary to consider how work performance varied during the periods when these changes were introduced. This is difficult because none of the reports of the Hawthorne studies provides actual figures covering the way in which output changed throughout Stage I. Consequently, one must work with such estimates as can be derived from the various graphs and charts of output-change that are supplied, and supplemented by occasional statements in the texts which give additional quantitative information.

AN EXAMINATION OF THE EVIDENCE: VARIATIONS IN SUPERVISORY PRACTICE AND VARIATIONS IN OUTPUT

For present purposes, Stage I may be divided into three phases: Phase I: the first three and a half months in the test room during which supervision seems to have been fairly consistently friendly, casual, and at low pressure; Phase II: a further interval of about seven months during which supervision became increasingly stern and close. This phase culminates in the dismissal of two of the five operators and their replacement by workers of rather special character and motivation. Phase III: a final long period during which output rose

⁷⁷ *Ibid.*, pp. 120–129; Roethlisberger and Dickson, *op. cit.*, pp. 63, 74, 86, 156, 167.

⁷⁸ Roethlisberger and Dickson, *op. cit.*, p. 162.

⁷⁹ Whitehead, *op. cit.*, vol. I, pp. 121–122. Roethlisberger and Dickson (*op. cit.*, pp. 60, 62) give no indication that the operators were paid for these hours not worked. Indeed, their account clearly implies that they were not so paid (*ibid.*, pp. 63–64). But Whitehead is quite explicit on this point.

⁸⁰ Roethlisberger and Dickson do report (*op. cit.*, p. 68) that the girls were paid for the half day on Saturdays which was not worked. They acknowledge that this "added a new factor to the situation which cannot be disregarded and which has to be taken into account in comparing this period with any other" (*ibid.*, p. 69). They take no further account of it, however, just as they take no further account of the unworked hours paid for on the occasions when the work day was shortened.

⁸¹ That is, up to the end of Experimental Period 7 in Roethlisberger and Dickson's terminology.

⁸² Roethlisberger and Dickson, *op. cit.*, pp. 58–59.

rapidly and there was a return to free and friendly supervision.

Supervision during Phase I. "Besides the girls who composed the group under study there was a person in the experimental room who was immediately in charge of the test." This was the test room observer whose two-fold function was "to keep accurate records . . . and to create and maintain a friendly atmosphere in the test room." He "assume(d) responsibility for most of the day to day supervision" while in other matters such as accounting, rate revision, and promotion, responsibility rested with the foreman.⁸³

It is quite clear from Roethlisberger and Dickson's account that during Phase I the supervisors did everything in their power to promote a free, cooperative, and non-coercive relationship.⁸⁴ At the outset of the study the girls "were asked to work along at a comfortable pace" and were assured "that no attempt would be made to force up production." They were led to expect changes in working conditions which might be "beneficial and desirable from the employees' point of view," and were told that there was no reason why "any (such) change resulting in greater satisfaction of employees" should not be maintained, and this "regardless of any change in production rate."⁸⁵ "The test room observer was chiefly concerned with creating a friendly relation with the operators which would ensure their cooperation. He was anxious to dispel any apprehensions they might have about the test and, in order to do this, he began to converse informally with them each day."⁸⁶ Some weeks after the study began, there was a friendly talk with the doctor about the physical examinations and ice cream was provided and a party planned. Also, the girls were "invited to the office of the superintendent who had talked to them, and in various other ways they had been made the object of considerable attention."⁸⁷ Although there had been from almost the beginning a good deal of talking among the

girls, a fairly permissive attitude had been taken about this.⁸⁸

Output during Phase I. There was "no appreciable change in output" on transfer to the test room,⁸⁹ but there was a "downward tendency" during the first five weeks thereafter,⁹⁰ despite facilities which "made the work slightly easier."⁹¹

At the end of five weeks, the new wage incentive system was introduced and output increased.⁹² From the output chart⁹³ this increase may be estimated at 4 or 5 percent. However, this increase must be accepted with some caution, for the investigators report that the "change in method of payment necessitated a change in piece-rates."⁹⁴ It was apparently judged that under the new conditions of work, (which did not include all of the types of relay assembled on the shop floor, and where there was one layout operator to five assemblers instead of one to six or seven as on the shop floor) new rates were necessary. We are told that "the chief consideration in setting the new piece rates was to determine a rate for each relay type which would pay the operators the same amount of money they had received in the regular department for an equivalent amount of work."⁹⁵ But it is well-established that the unreliability of time-study ratings can be expected to yield errors of at least 5 percent between different ratings of similar tasks.⁹⁶ So no great reliance can be placed on the observed 4 or 5 percent increase in output following the introduction of the new incentive system and the associated new piece-rates. Indeed, there is perhaps some recognition of this in Roethlisberger and Dickson's introductory comment that early in the study "a change in wage payment was introduced, a necessary step before the *experiment proper* could begin."⁹⁷ Phase I ends after fifteen weeks of friendly supervision with a some-

⁸³ *Ibid.*, pp. 22, 37.

⁸⁴ *Ibid.*, pp. 32-39.

⁸⁵ *Ibid.*, p. 33.

⁸⁶ *Ibid.*, p. 37.

⁸⁷ *Ibid.*, pp. 34, 39.

⁸⁸ *Ibid.*, p. 53.

⁸⁹ Pennock, *op. cit.*, pp. 301, 304.

⁹⁰ Roethlisberger and Dickson, *op. cit.*, p. 58.

⁹¹ *Ibid.*, pp. 33-34, 39.

⁹² *Ibid.*, p. 58.

⁹³ *Ibid.*, p. 56.

⁹⁴ *Ibid.*, p. 34.

⁹⁵ *Ibid.*, p. 35.

⁹⁶ Viteles, *Motivation . . .*, *op. cit.*, pp. 30-38.

⁹⁷ Roethlisberger and Dickson, *op. cit.*, p. 29, italics added.

what doubtful increase of 5 percent which occurred with the introduction of a preferred incentive system.

Supervision during Phase II. "The second phase . . . covering an interval of approximately seven months was concerned with the effects of various kinds of rest pauses."⁹⁸ The investigators emphasize that by the *beginning* of this phase not only was supervision friendly, but the relation between workers and supervisors was "free and easy."⁹⁹ Their account of actual supervisory behavior during succeeding months supports these claims. (i) On each of the four occasions when rest pauses were varied, the girls were consulted in advance, and on all but one occasion their expressed preferences were accepted. (ii) The investigators decided to pay the girls their bonuses monthly instead of weekly, but when the girls were told about this decision they objected and the plan was dropped. That the girls "felt free to express their attitudes" and that the investigators altered their plans out of regard for these attitudes is said to be "typical of the supervisory technique employed" which "proved to be a factor of utmost importance in interpreting the results of the study." (iii) Later the girls were given free lunches and were consulted about what should be served.¹⁰⁰

However, the problem of excessive talking among the girls worsened. No attempt had been made to prohibit talking, although four of the girls had been "given a talk regarding their behavior"¹⁰¹ Now this "lack of attention to work and preference for conversing together for considerable periods" was judged to be reaching such proportions that the "experiment was being jeopardized and something had to be done."¹⁰² A variety of disciplinary procedures of increasing severity were applied, but with little effect. Finally, the leaders in talking (operators 1A and 2A) were dismissed from the test room "for lack of cooperation which would have otherwise

necessitated greatly increased disciplinary measures."

Output during Phase II. There was no change in weekly output during this six-month period. "Total weekly output does not decline when rest pauses are introduced, but remains practically the same during all the rest period experiments."¹⁰³

Supervision during Phase III. At the beginning of Phase III,¹⁰⁴ the two dismissed girls were replaced by two girls chosen by the foreman. Something has already been said about the way in which these girls at once took and maintained the lead in output and about how one of them, who had a special need for more money took over the general leadership and discipline of the rest of the group. These points will bear underlining by direct quotation:

"When Operator 2 joined the group, her home was largely dependent upon her earnings, and within a few weeks her father lost his job and became temporarily unemployed. Thus, to her natural sense of responsibility was added the factor of poverty; and Operator 2 began to urge the remainder of the group to increase their output."¹⁰⁵

"Operators 1 and 2 were very definitely the fastest workers of the group in 1928, and this was freely recognized by the others."¹⁰⁶

"On the whole, from January to November 1928, the Relay Test Group showed no very marked developments apart from a growing tendency for the discipline to pass from the hands of the supervisor to those of the group itself, largely as represented in the person of Operator 2."¹⁰⁷

"Operator 2 became recognized as the leader of the group, both by the operators themselves and by the supervisor. It is doubtful whether any operator could have secured this position unless she had been the fastest worker, but the other qualifications possessed by Operator 2 were a high sense of the importance of the work for the group and a forceful personality."¹⁰⁸

"Op. 2. 'Oh! what's the matter with those other girls. I'll kill them.'"¹⁰⁹ (This expostu-

¹⁰³ *Ibid.*, p. 79.

¹⁰⁴ Actually on January 25, 1928, two days after the beginning of Phase III. Thus, the resulting sharp rise in output does not show fully on Roethlisberger and Dickson's weekly output charts (*op. cit.*, pp. 76, 78) until the second week of their Experimental Period 8.

¹⁰⁵ Whitehead, *op. cit.*, vol. I, pp. 122-123.

¹⁰⁶ *Ibid.*, p. 126.

¹⁰⁷ *Ibid.*, p. 124.

¹⁰⁸ *Ibid.*, p. 129.

¹⁰⁹ *Ibid.*, p. 127.

⁹⁸ *Ibid.*, p. 40. This phase actually extends from Aug. 8, 1927 to January 21, 1928, a period of twenty-four weeks.

⁹⁹ *Ibid.*, pp. 45-46.

¹⁰⁰ *Ibid.*, pp. 48-9, 51.

¹⁰¹ *Ibid.*, p. 38.

¹⁰² *Ibid.*, pp. 53-54.

lation was provoked by the output curves showing operators 3, 4, and 5 on a downward trend.)

From then on supervision again became increasingly friendly and relaxed. This friendliness of supervision often had a very tangible character. From the arrival of the new workers in the test room, the observer "granted them (all) more and more privileges." The preferred incentive system, the rest pauses, the free lunches, and the "parties" following the regular physical examinations all continued.¹¹⁰ In addition, within the next eight months the girls were first paid for half an hour per day not worked, and then for an hour a day not worked, and finally for Saturday mornings not worked. Approximately eight months after the arrival of the new girls, all these privileges except the preferred incentive system and the parties were withdrawn. The girls were warned in advance about this withdrawal of privileges and were assured that the new and heartily disliked conditions "would terminate after approximately three months." Despite this promise, the girls' work deteriorated immediately: they wasted time in various ways such as reading newspapers, eating candy, and going for drinks and the observer shortly "discovered that the girls were attempting to keep the output rate low . . . so as to make sure that rest pauses would be reinstated." The observer "again tried to stop the excessive talking" by "reprimand and threat." He told the girls that "unless excessive talking ceased it might become necessary to continue the experiment without rest pauses for a longer period."¹¹¹

At this point, the girls had been in the test room eighteen months and had achieved nearly all the eventual 30 percent increase in output. Yet it would seem that Operator 2, the incentive system, and the other privileges, as well as "reprimand and threat" played a significant part in determining the work behavior and output of the group. It is also clear from Roethlisberger and Dickson's account that for a great part of the time following the arrival of Operators 1

and 2, the girls worked very well and happily and that while they did so, supervision was relaxed and friendly and relations continued to be satisfactory. But there would seem to be good grounds for supposing that supervision became more friendly and relaxed because output increased rather than vice versa.

Output during Phase III. Output for the whole group rose markedly during the several months after the dismissal of 1A and 2A, owing chiefly to the contributions from the new operators.¹¹² Thereafter, the group's total output rose more slowly for a further year (with a temporary drop when the Saturday morning shift was discontinued for a time).

SUMMARY OF EVIDENCE ABOUT SUPERVISION AND OUTPUT

(i) Apart from a doubtful 4–5 percent increase following the introduction of a preferred incentive system, there was no increase in weekly output during the first nine months in the test room, despite a great deal of preoccupation on the part of the supervisors with friendliness towards the workers, with consultation, and the provision of a variety of privileges not enjoyed on the factory floor.

(ii) From the beginning of what Roethlisberger and Dickson describe as the "experiment proper," that is, after the period in which the new incentive system was introduced, there was no increase in weekly output during the next six months. When it became apparent that free and friendly supervision was not getting results, discipline was tightened, culminating in the dismissal of two of the five girls.

(iii) The dismissed girls were replaced by two girls of a special motivation and character who *immediately* led the rest in a sustained acceleration of output. One of these girls who had a special need for extra money rapidly adopted and maintained a strong disciplinary role with respect to the rest of the group. The two new girls led the way in increased output from their arrival till the end of the study.

(iv) Total output per week showed a sig-

¹¹⁰ Roethlisberger and Dickson, *op. cit.*, pp. 71, 72, 77.

¹¹¹ *Ibid.*, pp. 70–72.

¹¹² *Ibid.*, Figure 7, p. 78.

nificant and sustained increase only after the two girls who had the lowest output¹¹³ were dismissed and replaced by selected output leaders who account for the major part of the groups' increase, both in output rate and in total output, over the next seventeen months of the study.

(v) After the arrival of the new girls and the associated increase in output, *official* supervision became friendly and relaxed once more. The investigators, however, provide no evidence that output increased because supervision became more friendly rather than vice versa. In any case, friendly supervision took a very tangible turn by paying the girls for time not worked the piece-rate was in effect increased.

DISCUSSION AND CONCLUSIONS

The critical examination attempted here by no means exhausts the gross error and the incompetence in the understanding and use of the scientific method which permeate the Hawthorne studies from beginning to end. Three further studies were conducted: the Bank Wiring Observation Room Study; the Interviewing Program; and the Counselling Program. These studies cannot be discussed here, but I believe them to be nearly as worthless scientifically as the studies which have been discussed.¹¹⁴ This should not be surprising, for they arose out of "evidence" found and conclusions reached in the earlier studies and were guided by and interpreted in the light of the strongest preconceptions based on the conclusions of the earlier studies.

There are major deficiencies in Stages I, II and III which have hardly been touched on: (i) There was no attempt to establish sample groups representative of any larger

population than the groups themselves. Therefore, no generalization is legitimate. (ii) There was no attempt to employ control data from the output records of the girls who were *not* put under special experimental conditions. (iii) Even if both of these points had been met, the experiments would still have been of only minor scientific value since a group of five subjects is too small to yield statistically reliable results. Waiving all these points, it is clear that the objective evidence obtained from Stages I, II, and III does not support any of the conclusions derived by the Hawthorne investigators. The results of these studies, far from supporting the various components of the "human relations approach," are surprisingly consistent with a rather old-world view about the value of monetary incentives, driving leadership, and discipline. It is only by massive and relentless reinterpretation that the evidence is made to yield contrary conclusions. To make these points is not to claim that the Hawthorne studies can provide serious support for any such old-world view. The limitations of the Hawthorne studies clearly render them incapable of yielding serious support for any sort of generalization whatever.

If the assessment of the Hawthorne studies offered here is cogent, it raises some questions of importance for university teachers, especially for teachers concerned with courses on industrial organization and management. How is it that nearly all authors of textbooks who have drawn material from the Hawthorne studies have failed to recognize the vast discrepancy between evidence and conclusions in those studies, have frequently misdescribed the actual observations and occurrences in a way that brings the evidence into line with the conclusions, and have done this even when such authors based their whole outlook and orientation on the conclusions reached by the Hawthorne investigators? Exploration of these questions would provide salutary insight into aspects of the sociology of social scientists.

¹¹³ *Ibid.*, p. 162.

¹¹⁴ For substantiation of this judgment with respect to the Bank Wiring Observation Room Study see A. J. Sykes, "Economic Interest and the Hawthorne Researches: A Comment," *Human Relations*, 18 (August, 1965), pp. 253-263.