AMERICAN SOCIOLOGICAL REVIEW
October 1978 Volume 43 Number 5

THE HAWTHORNE EXPERIMENTS: FIRST STATISTICAL INTERPRETATION*

RICHARD HERBERT FRANKE
Worcester Polytechnic Institute

JAMES D. KAUL
Environmental Protection Agency

American Sociological Review 1978, Vol. 43 (October):623–643

A guide is provided to the proceedings of the Hawthorne experiments, and experimental data are now made readily available. Data from the main experiment (that in the first relay assembly test room at Western Electric) are interpreted statistically for the first time. Quantitative analysis of this quasi experiment is accomplished by time-series multiple regression using nearly five years of data. This analysis demonstrates that experimental variables account for some 90% of the variance in quantity and quality of output, both for the group and for individual workers. Imposition of managerial discipline, economic adversity, and quality of raw materials provide most explanation, obviating the need to draw upon less clearly definable human relations mechanisms. For decades the Hawthorne studies have provided a rationale for humane approaches in the organization of work by suggesting that considerate or participative treatment of workers led to better economic performance. The present analysis suggests, to the contrary, that humanitarian procedures must provide their own justification.

The massive Hawthorne experiments of some 50 years ago serve as the paradigmatic foundation of the social science of work.[1] The insights gleaned from these experiments provide a basis for most current studies in human relations as well as for subareas such as participation, organizational development, leadership, motivation, and even organizational design. But aside from visual inspection and anecdotal comment,[2] the complex of data obtained during the eight years of the Hawthorne experiments has never been subjected to thorough-going scientific analysis. Indeed, as was pointed out in this journal by Carey (1967), the data necessary for statistical analysis are not available in the scientific literature. It is the purpose of this report to make the Hawthorne data accessible, to interpret systematically the most important of these, and to draw from the results thus obtained some conclusions regarding the use of social science in industry.

Since interpretation and criticism of the Hawthorne studies to date have been little

* Address all communications to: Richard H. Franke; Department of Management; Worcester Polytechnic Institute, Worcester, MA 01609.

We are indebted to the Western Electric Corporation, to the Library and the School of Business Administration of the University of Wisconsin, Milwaukee, and to the Worcester Polytechnic Institute for facilitating and supporting this study. For their help in obtaining the Hawthorne documents, we thank particularly E. L. Byrom and Florence Hybl of Western Electric; William Roselle, Stanley Mallach, and Steven Smith of UWM Library; and Ruth Zubrensky and George Carian. We also thank two anonymous reviewers and the following persons for their comments on earlier drafts: Bernard Bass, Robert Dubin, Elke Franke, Arthur Gerstenfeld, Donald Gerwin, Chadwick Haberstroh, Alex Inkeles, Wilbert Moore, Stephanie Pluskota, James Price, and Winston Ring. An earlier report of these findings was presented at the annual meeting of the American Association for the Advancement of Science, Washington, 1978.

[1] For a selection of studies testifying to the importance of the Hawthorne studies in the development of applied social science, cf. Homans (1941; 1950), Friedman (1946), Miller and Form (1951), Viteles (1953), Blum and Naylor (1968), Sanford (1973), Cass and Zimmer (1975), and Locke (1976).

[2] Such interpretation has been taken to considerable lengths, as in the work of Roethlisberger and Dickson (1939), Homans (1950), and Whitehead (1938). The latter author did also employ statistical procedures, but without application to the major dependent variables in the various experiments.

more than opinion, most of this introduction will be a simple description of the Hawthorne experiments over 1924 to 1933, with brief note of the conclusions and impact of these studies.[3] Systematic review of the secondary literature is presented following the analytical section, so that these evaluations may be judged in light of the results of quantitative analysis.

The Hawthorne studies began in 1924 at the Hawthorne plant of the Western Electric Company in Chicago with an inquiry by the National Academy of Sciences and Western Electric into relationships between illumination levels and worker production rates. Inexplicably worker output and job satisfaction generally increased regardless of increase or decrease in illumination. Their curiosities piqued, Western Electric management and social scientists from the Harvard School of Business Administration initiated experiments to examine effects of social as well as physical factors upon work efficiency. A chronology of the experiments is presented in Figure 1. The exploratory illumination experiments (1924–27) were followed by the main Hawthorne experiment, in the first relay assembly test room (1927–33), and by four derivative experiments (1928–32). The first four experimental programs were reactive; that is, conditions were manipulated by the experimenters, who then noted changes in work satisfaction and performance. The final two experiments did not include advertent manipulation of independent variables. However, the presence of interviewers and observers was itself a change in the conditions of work.

A flowchart and description of events is presented in Figure 2.[4] From sole attention to environmental conditions of work in (1) the illumination experiments, the studies expanded in (2) the first relay experiment to scrutinize effects of work environment, physical requirements, management, and social relations upon output. All issues dealt with subsequently were initiated, at least broadly, in the first relay experiment. The derivative studies were: (3) the second relay experiment, which tested and discounted effects of small group incentive payment; (4) the mica splitting experiment, which tested and discounted effects of rest pauses upon performance; (5) the interviewing program, which indicated that relations with management and with peers were important to worker satisfaction, and that informal group organization could be used by workers to regulate and reduce the pace of their work; (6) the bank wiring observation, which confirmed the latter conclusion regarding output restriction, and thus underlined the importance of social relations among workers. Counseling, supervisory training, and other nonexperimental programs also were undertaken by Western Electric to make use of the conclusions from the six experiments. In experiments (2), (3), (4), and (6), research attention focused on small group activities. Three separate groups of five female workers each were involved in the first relay, second relay, and mica splitting experiments, while 14 male workers participated in the bank wiring study. The (1) illumination and (5) interviewing studies, on the other hand, involved whole departments of workers.

The researchers concluded from both the primary and the derivative experiments that measured experimental variables had little effect, but that the unmeasured quality of human relations of workers to management and peer group was responsible for most output improvements observed in the first four experiments. This rather unspecific conclusion, providing a foundation for modern humanitarian and human relations approaches to work, led other researchers to focus upon worker satisfaction, as in the Ohio State supervision studies and their descendants (cf. Fleishman et al., 1955; Miner, 1965), to studies of authoritarianism (cf. Sales, 1966; Vroom, 1960), informal organization (Whyte, 1955), leadership (Bass, 1960; Stogdill, 1974), participative management (cf. Likert, 1967; Marrow, 1975; Pusić,

[3] The primary sources describing and interpreting the Hawthorne experiments are the monumental *Management and the Worker* of Roethlisberger and Dickson (1939), two volumes of graphs and description by Whitehead (1938), and description and social application by Elton Mayo (1933; 1945; 1947).

[4] The description and conclusion are extracted from Roethlisberger and Dickson (1939).

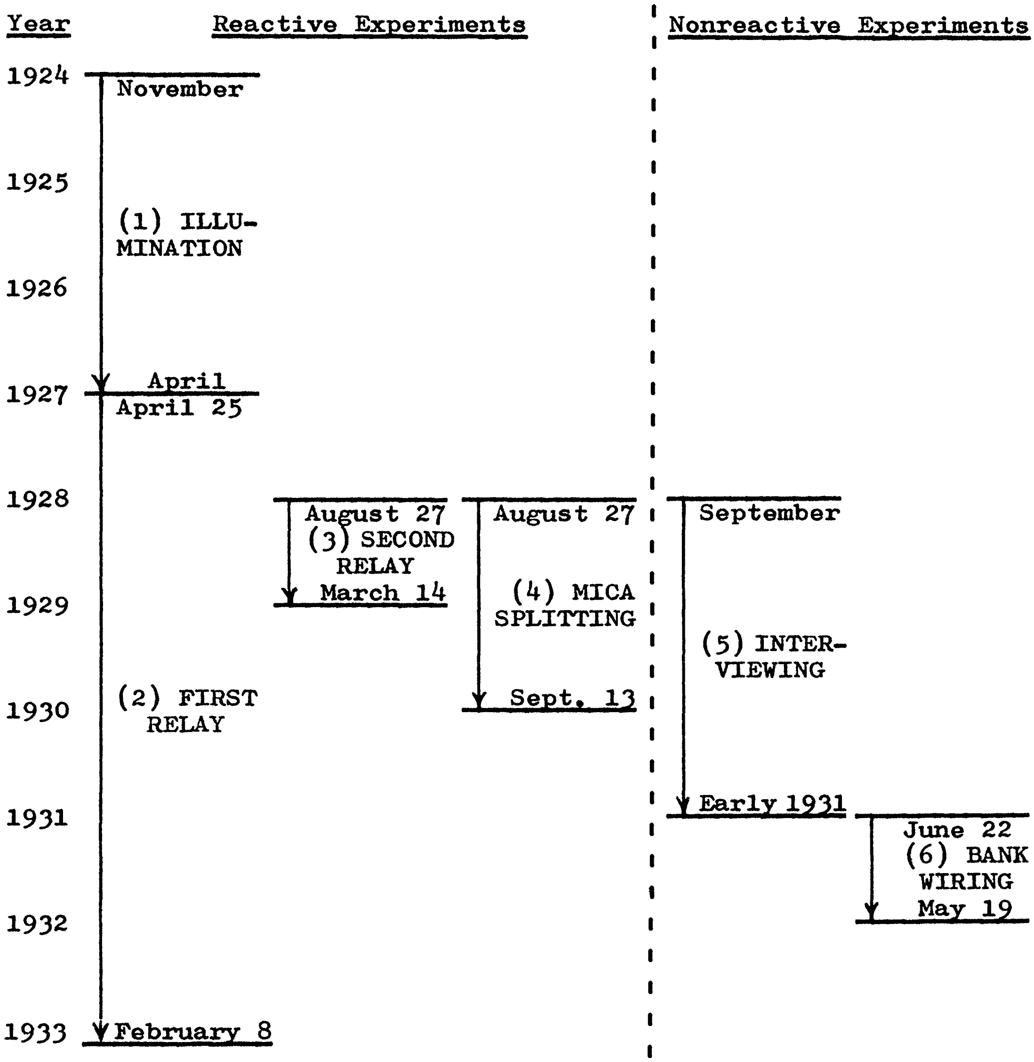

Figure 1. Chronology of the Hawthorne Experiments

1973), and to the use of sensitivity training and related techniques in organizational development (cf. Beer, 1976; Bradford et al., 1964; Golembiewski and Blumberg, 1970). It should be stated here that the initial concern of the Hawthorne experiments was with output. Concentration upon worker satisfaction in subsequent studies is sometimes justified by assuming it to be an intervening factor in job performance. This is an assumption often made in practice if not in word (as by Kahn, 1975, and Price, 1968), in the face of contradicting evidence (cf. Locke, 1976; Vroom, 1964). Attempts to demonstrate empirically a linkage between human relations and work performance have not received primary attention since the time of the seminal Hawthorne experiments. To cast further doubt upon the human relations conclusion, we should again note that there have never been meaningful statistical analyses of the data from the Hawthorne experiments. The absence of statistical analysis may have resulted from the nature of the experiments: (a) there were no control groups other than the experimental groups themselves prior to manipulation; (b) since these were field experiments over extended periods

Flowchart	Experimental Variables Manipulated or Observed: Physical Work Environment	Physical Work Requirements	Management and Supervision	Social Relations of Workers	Description and Conclusion
(1) ILLUMINATION	X				Three exploratory studies that suggested human factors rather than physical working conditions determined worker satisfaction and performance.
(2) FIRST RELAY	X	X	X	X	The major Hawthorne experiment, testing effects on performance of rest pauses, shorter work periods, and increased worker autonomy and of small group incentive pay. This study concluded that benefits to worker performance resulted from improved human relations, and to a lesser extent from rest pauses.
(3) SECOND RELAY			X		Derivative experiment suggesting only moderate effects of small group incentive pay upon worker performance.
(4) MICA SPLITTING		X			Derivative experiment suggesting only moderate effects of rest and shorter work periods upon worker performance.
(5) INTERVIEWING			X	X	Derivative survey reinforcing prior conclusions regarding importance of social interactions (worker-worker and worker-supervisor) in the satisfaction of workers. First indications, during intensive interviews early in 1931, of problems resulting from employee interrelations, especially in restriction of output.
(6) BANK WIRING				X	Derivative observations noting the effectiveness of social interactions in a large group of workers in standardizing the pace of work (restricting output during period of economic depression).

Figure 2. Flow, Content, and Conclusions of the Hawthorne Studies

of time, it would have been difficult to eliminate all extraneous variables from the experiment.[5]

Fortunately, quasi-experimental approaches and time-series analytical procedures have been developed since the original researchers' interpretations of the Hawthorne experiments (Mayo, 1933; Roethlisberger and Dickson, 1939; Whitehead, 1938). These methods allow the use of periodic data, with testing for effects of measured variables and adjustment for changes over time in unmeasured historical factors. Rough schematic approaches have been suggested for analysis in education and psychology by Campbell and Stanley (1963) and Cook and Campbell (1976), while quantitative regression procedures have been developed in the field of econometrics (cf. the analytical procedures described below).

One investigation at Hawthorne, the first relay experiment, included a variety of dependent and independent variables that could be expressed quantitatively over 23 experimental periods,[6] allowing convenient use of time-series regression. Measures of quantity and quality of output could be obtained for the group of five workers and for each individual, and measures of independent variables also could be obtained. Rest pauses, hours of work per day, and days of work per week were intentionally manipulated, and a small group incentive system was introduced. A number of inadvertent categoric changes which occurred over the five years of the experiment also could be identified. Most interesting of these changes were the replacement by management of two of the workers after period 7, because of their unsatisfactory attitudes in response to requests for greater diligence and more output, and the onslaught of the great depression early in period 15. The analyses that follow will use the available evidence to test directly for the sources of differences in worker performance over time, to determine whether the substantial performance variances obtained in the experiment can be explained quantitatively.

[5] The experimental and analytical approaches of the Hawthorne experimenters have been criticized in these and other lights by numerous authors, including Carey (1967), Cook and Campbell (1976), and Farris (1969), and by those reviewed by Landsberger (1958).

[6] There was also a 24th period, from March 1, 1932, to February 8, 1933, which is not considered in the present analyses. During this final period, all five operators were laid off and replaced by more senior workers who were inexperienced in the assembly of relays.

METHOD

Data

The original documents from the Hawthorne experiments were reviewed and then borrowed during visits to the Hawthorne plant between November 1976 and May 1977. Copies of the documents are now on microfilm in the libraries of the University of Wisconsin, Milwaukee, and the Worcester Polytechnic Institute. The data of the first relay experiment are summarized in Appendix 1 (group) and Appendix 2 (individuals). Quantity of output is recorded as net hourly and net weekly rate per worker, while quality of output is recorded as repair time required per day. Hours worked per day and per week and the number of weeks per experimental period describe the basic work schedules, with time taken for scheduled and voluntary rest pauses reducing actual work time. Categoric changes in working conditions are expressed as dummy variables of zero to one for managerial discipline (the replacement by management of two of the five workers, with one of the replacements assuming the role of straw boss), and for the occurrence of the economic depression, the supply of defective raw materials for two periods, the temporary voluntary replacement of one worker, and for the change from a large group to a small group incentive system of pay after the first two experimental periods. A list of these variables and their dimensions is provided in the first results table presented below. Data were available over the 23 periods for all variables, except for repair time (18 periods) and voluntary rest time (21 periods), each unmeasured in several early and late periods.

Analytical Procedures

The data of the first relay experiment are suitable for rigorous analysis using time-series econometric techniques. With these, even influences of inadvertent experimental changes can be examined specifically.[7] In addition, influence of other (generalized) historical factors as well as the passage of time can be measured as *serial correlation* (Durbin and Watson, 1950; 1951), and any effects upon the statistical independence of sequential sets of data can be removed (using the approach of Theil and Nagar, 1961, as described by Johnston, 1963, and Elliott, 1973).

Analyses are directed toward explanation of differences in output over time for the group, as well as for individual workers. As a first step, zero-order correlations are examined for the first relay group—for the entire 23 periods, and separately for the seven periods prior to replacement of two unsatisfactory workers and for the 16 periods after this exercise of managerial discipline. In the second step, the best multiple regressions are determined for the group using all available periods for each of the three production measures, with correction for serial correlation where necessary. As a third and final step, this procedure is repeated for the production rates of each of the individual workers, first by forcing the group models upon the individual data and then by determining whether alternate models provide greater variance explanation.[8]

RESULTS

Zero-Order Relationships for the Group

Results from correlating all group data over the experimental periods are presented in Table 1, employing the maximum number of periods available for each pair of variables. The group dependent variables are net measures of output quantity per worker per hour, and per week, as well as a negative measure of output quality which also is subsumed in the quantitative measures (repair time required per worker per day). These three performance variables were correlated over periods with each other and with the 12 independent experimental variables in work schedules, rest pauses, and categoric changes of working conditions that occurred during the five years analyzed for the first relay experiment. Results are presented separately for the entire 23 periods, for the first seven periods with the original five workers, and for the final 16 periods with the group containing three original and two replacement workers.

Because the variables are measured for sequential periods of time and their generalized historical dependence or serial correlation has not yet been measured, it would be premature to evaluate relationships in terms of statistical significance. However, there are in some cases very substantial simple relationships between group performance and experimental variables, yielding high levels of variance explanation (R^2). For example, period-by-period differences in rate of hourly output over 23 periods can be explained in large part (25% or more) by the categoric variables of managerial discipline and economic depression, with group hourly production apparently improved by management's replacement of two workers and by the depression. Also, fewer net hours per week, more scheduled rest time, and use of small group incentives appear to have improved hourly production rates.[9] Since most intercorrelations of

[7] For a treatment of regression equations containing dummy variables, cf. Johnston (1963).

[8] Zero-order correlation coefficients also have been calculated for individuals over the 23, 7, and 16 periods, and group and individual regression equations have been calculated for the first 7 and for the remaining 16 periods. No additional findings of note were obtained in these calculations, which thus are not included in the present report.

[9] The effects of incentive system and rest pauses upon output also were shown in the data of the second relay and mica splitting experiments. (The data are presented but not analyzed by Roethlisberger and Dickson, 1939: 132, 148.) The incentive effect was tested in the second relay experiment, where a group of five operators worked in one period with the existing large group incentive arrangement, in a second period with the small group of five as basis for incentive pay, and in a final period after return to the large group incentive system. The mean rates of production per worker rose from 1,634 to 1,840 and then back to 1,531 unit components per hour. With the earlier periods serving as controls for the same workers in the next periods, t-test analysis shows a significant difference only between periods 1 and 2 ($t = 2.54$, $p < .05$). That is, there was a significant 12.6% improvement in rate of production

Table 1. Experimental Correlations—Group Outputs and Repair[a]

	Periods 1–23			Periods 1–7			Periods 8–23		
	Hourly Output	Weekly Output	Repair Time	Hourly Output	Weekly Output	Repair Time	Hourly Output	Weekly Output	Repair Time
Dependent Variables									
(1-G) Hourly Output, units/hr.									
(2-G) Weekly Output, units/wk.	.197			.973			−.356		
(3-G) Repair Time, min./day	.225	.103		.693	.524		.187	.036	
Independent Variables									
Work Schedule									
(4) Hours per day	−.656	.378	.117	—[d]	—	—	−.549	.702	.199
(5) Days per week	−.629	.585	−.195	—	—	—	−.624	.891	−.174
(6) Net Hours per week[b]	−.758	.489	−.155	−.964	−.878	−.778	−.691	.920	−.077
(7) Weeks per period	.394	.480	.037	.349	.426	−.661	.040	.394	.028
Rest Pauses									
(8) Scheduled Rest Stops, no./day	.261	.036	.374	.761	.626	.884	.425	−.229	.324
(9) Scheduled Rest Time, min./day	.693	.161	.427	.964	.877	.788	.425	−.229	.324
(10-G) Voluntary Rest Time, min./day	−.466	−.096	−.356	−.612	−.402	−.852	−.256	.129	−.189
Categoric Changes in Working Conditions[c]									
(11) Managerial Discipline	.887	.343	.118	—	—	—	—	—	—
(12) Economic Depression	.791	−.096	.256	—	—	—	.917	−.351	.245
(13) Defective Raw Materials	.101	.275	.806	—	—	—	−.198	.226	.863
(14) Temporary Replacement of Oper. 5	.244	.594	.651	—	—	—	.064	.567	.694
(15) Small Group Incentive	.552	.272	—	.771	.840	—	—	—	—

[a] All data employed are per-capita averages over all periods for which data were available (see Appendix 1); these Pearson product-moment correlations are for periods common to both variables indicated (in total 18, 21, or 23), and are uncorrected for serial correlation. [b] Hours per day times days per week, minus scheduled rest time. [c] Change indicated=1, otherwise=0. [d] Dashes indicate no data or no variance in a variable.

these independent variables are smaller than their correlations with the dependent variable, prospects appear good for multivariate explanation of a large portion of variance in hourly output by known experimental variables. Separate correlations are presented in Table 1 for group hourly output with independent variables over periods 1–7 and 8–23—before and after the imposition of managerial discipline expressed by the replacement of operators 1A and 2A with operators 1 and 2. During the early periods, more rest time (leading to fewer net hours) and use of small group incentive payment seem to have been beneficial. In later periods, the economic depression and fewer net hours appear to have benefited hourly output. These shorter-term results are consistent with correlations over the entire 23 periods, where, in addition, the apparent effect of managerial discipline is shown.

Correlations for weekly output rates are also presented in Table 1, but show strong simple relationships over 23 periods only with the temporary replacement of operator 5 and with the days (and net hours) worked per week. A greater number of hours worked per week appears to have offset lower hourly rates when viewing weekly output. During periods 1–7 more rest time (reflected in slightly fewer net hours) and the introduction of small group incentives seem to have aided weekly output, while during periods 8–23 more net hours and the replacement of operator 5 seem to have been beneficial. In sum, the replacement of operator 5, more hours per week, more rest time, and use of small group incentives seem to have benefited group weekly output. Neither managerial discipline nor the economic depression shows the strong simple relationship to weekly output seen in Table 1 for hourly output. The replacement of operator 5 seems to have been important to weekly but not to hourly output, and more net hours seem to have benefited weekly output while detracting from hourly rate of output. On the zero-order surface, only rest time and small group incentive payment seem to have been useful to both rates of output.

Correlations for the third performance variable of repair time required per day (extent of poor quality output) also are presented in Table 1. Poor quality is strongly related to the use of defective raw materials and to the temporary replacement of operator 5, for periods 3–20 and for the later periods (8–20). During the early periods (3–7), lower quality is associated with more scheduled rest stops (and the attendant fewer net hours and less voluntary rest time taken) and with shorter experimental periods. Thus poorer quality output appears to result from a combination of factors which might affect quality more or less mechanically—from poor raw materials, an inexperienced worker, and from breaks in work routine by more frequent daily rest stops and more frequent changes in work schedule and conditions.

Multiple Analysis: Group Regression Equations

Each of the three dependent variables for the group has been regressed stepwise upon the 12 independent variables identified in Table 1 (data in Appendix 1). The results are presented in Table 2 as regression equations. Also presented are the multiple coefficients of determination (R^2 and cR^2, the latter "corrected" for the number of independent variables), the Durbin-Watson coefficients of serial correlation (DW), stepwise variance explanations, and the regression equations after correction for serial correlation, if necessary.

Differences in rates of hourly output by the first relay group are explained in model 1 through managerial discipline (79%), economic depression (an additional 14%), and through scheduled rest time (4%). Most of this 97% variance explanation appears to have resulted from the imposition of managerial discipline, which

---

that appeared to result from the use of a small group incentive system. Similarly for the five workers in the mica splitting test, analysis indicates a significant and even more substantial (15.5%) improvement in hourly production rate, apparently resulting from the reduction of fatigue by use of rest pauses and fewer working hours ($t = 3.34$, $p < .02$). Other positive effects of performance-contingent incentives and of fatigue reduction upon output rates have been reported in various settings (cf. Bass and Barrett, 1972; Cherrington et al., 1971; Taylor, 1911).

Table 2. Stepwise Regression Equations—Group [a]

	Hourly Output	Intercept	Managerial Discipline	Economic Depression	Scheduled Rest Time
Model 1:	X_{1-G} =	50.50	+ 8.64 X_{11}	+ 6.18 X_{12}	+ 0.18 X_9
			$r_p = .928$	$r_p = .894$	$r_p = .758$
$R^2 = 97.10\%$, $cR^2 = 96.64\%$, DW = 1.58 ($p < .05$)					
Stepwise variance explan.: [b]			78.72%	+ 14.46%	+ 3.91%
Corrected for serial correlation (factor = 0.25):					
	X_{1-G} =	38.29	+ 8.26 X_{11}	+ 6.34 X_{12}	+ 0.16 X_9
			$r_p = .894$	$r_p = .859$	$r_p = .692$
			($p = .0000$)	($p = .0000$)	($p = .0007$)
$R^2 = 94.48\%$, $cR^2 = 93.56\%$ ($p = .0000$), DW = 1.95 (NS)					

	Weekly Output	Intercept	Net Hours Per Week	Managerial Discipline	Economic Depression	Scheduled Rest Time	Small Group Incentive
Model 2: [c]	X_{2-G} =	− 716.33	+ 66.52 X_6	+ 393.78 X_{11}	+ 245.97 X_{12}	+ 7.67 X_9	+ 105.32 X_{15}
			$r_p = .977$	$r_p = .942$	$r_p = .845$	$r_p = .731$	$r_p = .423$
			($p = .0000$)	($p = .0000$)	($p = .0000$)	($p = .0004$)	($p = .0713$)
$R^2 = 96.57\%$, $cR^2 = 95.56\%$ ($p = .0000$), DW = 1.98 (NS)							
Stepwise variance explan:			23.88%	+ 56.48%	+ 7.50%	+ 7.97%	+ 0.75%

	Repair Time	Intercept	Defective Raw Mtls.	Scheduled Rest Stops	Economic Depression	Weeks per Period
Model 3:	X_{3-G} =	17.88	+ 25.33 X_{13}	+ 2.58 X_8	+ 7.50 X_{12}	− 0.35 X_7
			$r_p = .951$	$r_p = .763$	$r_p = .770$	$r_p = -.724$
$R^2 = 92.38\%$, $cR^2 = 90.03\%$, DW = 2.32 ($p < .05$)						
Stepwise variance explan.: [b]			64.98%	+ 14.01%	+ 5.00%	+ 8.38%
Corrected for serial correlation (factor = − 0.40):						
	X_{3-G} =	22.36	+ 25.46 X_{13}	+ 3.08 X_8	+ 6.97 X_{12}	− 0.27 X_7
			$r_p = .964$	$r_p = .808$	$r_p = .821$	$r_p = -.594$
			($p = .0000$)	($p = .0005$)	($p = .0003$)	($p = .0251$)
$R^2 = 94.90\%$, $cR^2 = 93.21\%$ ($p = .0000$), DW = 2.07 (NS)						

[a] Independent variables are presented in order of appearance in stepwise multiple regression; n = 23 periods for output models and n = 18 periods for repair model.
[b] Stepwise variance explanation prior to correction for serial correlation.
[c] Stepwise model independent variables included net hours per week, excluding the highly correlated hours per day and days per week (cf. fn. 10).

included better performing replacement workers as well as the disciplinary example, from the beginning of period 8.

However, the time-series data used to obtain model 1 are not statistically independent, as indicated by the Durbin-Watson coefficient obtained (which is significantly distant from the neutral point of 2.00; cf. Durbin and Watson, 1951: Table 5). That is, the passage of time and unspecified historical factors appear to have influenced the regression residuals, making the use of critical values tables questionable in testing for significance of the regression equation. Correction for serial correlation using the Theil-Nagar (1961) approach is successful, as shown by the resulting nonsignificant Durbin-Watson coefficient of serial correlation. Variance explanation after correction is 94%, and the three independent variables show substantial and highly significant partial correlations with hourly output. Slope coefficients are altered little by correction: managerial discipline apparently resulted in a production increase of eight or nine units per worker per hour, the depression in an increase of six units, and scheduled rest time which ranged from zero to 30 minutes per day resulted in an increase of up to five units per worker per hour. As shown in Appendix 1, hourly output rose from some 50 units per worker per hour up to nearly 72 units by the end of the experiment, and most of this change is accounted for through the above slope coefficients.

Regression of weekly output of the first relay group upon experimental variables in model 2 yields as explanatory variables net hours per week,[10] then managerial discipline, economic depression, scheduled rest time, and finally the introduction of small group incentive payment. Although hours per week enters first as a regression variable, it provides only 24% variance explanation. Managerial discipline is again the major explanatory variable (56%), followed by the depression and rest time (8% each) and by small group incentive (1%), for total variance explanation of 97%. Since there is no serial correlation, the equation and each partial relationship can be tested for significance. All are substantial and highly significant, except for small group incentive ($.05 < p < .10$), which did not enter the stepwise regression for hourly output of the group. Slope coefficients in the weekly equation are comparable with those for hourly output (considering that the average work week contained 42 net hours), with additional factors of hours worked and the introduction of small group incentives. As shown in Appendix 1, weekly output ranged from about 2,100 to 3,200 units per worker, averaging some 2,600 units. Changes in net hours worked, with a range from 30.33 to 48 hours per week, could have accounted for a range of 1,175 units of weekly output per worker if not offset by changes in other variables. Managerial discipline and the economic depression could account for 394 and 246 units per week, and small group incentives for 105 units per week.

The third regression model presented in Table 2 seeks explanation of differences in quality of production as measured by repair time per worker per day. Stepwise regression for the first relay group shows explanation by the use of defective raw materials (65%), more frequent scheduled rest stops (14%), the economic depression (5%), and by fewer weeks per experimental period (8%). Total variance explanation is 92%. The acknowledged provision of defective raw materials in periods 14 and 15 appears to be the primary source of quality difficulty during the experiment, followed by increased disturbance of work routine by more rest stops and by shorter work periods containing unchanged working conditions, all apparently aggravated through stress induced by the depression. Correction of serial correlation, over the 18 periods for which repair data were available, did not substantially change slope coefficients. Interpretation of regression slopes shows the supply of defective raw materials accounting for 25 minutes of repair time per day, scheduled rest stops (zero to six) for up to 18 minutes, the depression for seven minutes, and shorter experimental periods (two to 31 weeks) for a range of about eight minutes of repair time per day.

For each group dependent variable in Table 2, there are measured experimental variables which explain well over 90% of the variance observed in production characteristics. Most of the difference in quantitative production of the group is explained by the replacement of two mediocre workers by others who from the outset demonstrated better performance. Most of the difference in quality of production is explained by difference in quality of raw materials. However, for an understanding of the meaning of these statistical results, replications are required at the level of individual workers—to measure for workers 1A plus 1 and 2A plus 2 the effects of replacement, and to measure the effects of this disciplinary example upon workers 3, 4, and 5.

Replication: Individual Multiple Regressions[11]

In the remaining tables, multiple regression results are presented interpreting the individual performance data in Appendix 2. The regressions include initially only the independent variables found useful in

[10] In this as in all other models presented, the variable net hours was included for potential regression equations, but the hours per day and days per week which are its constituents were excluded. However, models which included these variables also were examined. In no case did the resulting equation possess variance explanation superior to that of the corresponding model presented in this report.

[11] These analyses at a lower level of aggregation may be viewed as testing to guard against the ecological fallacy in making inferences based upon group data (cf. Blalock, 1961; Dogan and Rokkan, 1969; Galtung, 1967; Robinson, 1950; Thorndike, 1939).

group models (Table 2). Wherever independent stepwise multiple regressions yield results different from the group equations, these results are presented in the table footnotes. All results presented have been corrected for serial correlation (except for two instances of incomplete correction), and the variance explanations presented have been adjusted to allow for number of independent variables.

Multiple regressions for individual rates of hourly output are presented in Table 3. For workers 1, 2, 3, and 4, all three independent variables of group model 1 are significantly related to hourly output. Independent stepwise multiple regression shows no further variables to be important for workers 1, 2, and 4, but the additional variable of small group incentive payment benefiting worker 3. When model 1 is applied to worker 5, only the econimic depression relates significantly to hourly output. Independent stepwise multiple regression for worker 5 does show managerial discipline as a positive factor, with the additional and negatively related variables of defective raw materials (X_{13}) and the number of hours worked per week (X_6). The best individual equations from Table 3 provide explanation of variance in hourly output which ranges from 96% for operators 2A and 2 down to 66% for operator 4. Managerial discipline appears to have had the greatest effect upon those most directly involved in the replacement (operators 1A plus 1 and 2A plus 2), but also is significant for operators 3, 4, and 5, with smaller slope coefficients. Depression slope coefficients are about the same for operators 1, 2, 4, and 5, and lower but still significant for operator 3. Scheduled rest time is about equally important to the hourly output of operators 1, 2, 3, and 4, but unimportant to operator 5 (for whom fewer working hours seem to have been useful). Defective raw materials appear to have had an adverse effect upon the hourly output of operator 5 only.

Model 2 results for individual weekly output rates, presented in Table 4, are

Table 3. Replication of Regression Equation 1—Individual Hourly Output (X_1) upon Independent Variables from Table 2

X_1 for Operator(s)	Intercept	Managerial Discipline	Economic Depression	Scheduled Rest Time	Variance Explanation	Durbin-Watson Coefficient
1A + 1[a]	29.67	+ 10.84 X_{11} $r_p = .840$ ($p = .0000$)	+ 7.59 X_{12} $r_p = .758$ ($p = .0001$)	+ 0.19 X_9 $r_p = .583$ ($p = .0070$)	87.15%	1.67
2A + 2A[b]	49.60	+ 13.65 X_{11} $r_p = .939$ ($p = .0000$)	+ 8.07 X_{12} $r_p = .874$ ($p = .0000$)	+ 0.18 X_9 $r_p = .641$ ($p = .0017$)	96.42%	1.67
3[b]	51.63	+ 6.35 X_{11} $r_p = .835$ ($p = .0000$)	+ 2.74 X_{12} $r_p = .591$ ($p = .0048$)	+ 0.21 X_9 $r_p = .749$ ($p = .0001$)	91.12%	1.74
4[c]	21.78	+ 6.55 X_{11} $r_p = .637$	+ 5.78 X_{12} $r_p = .595$	+ 0.17 X_9 $r_p = .570$	65.79%	1.33
5[d]	20.83	+ 2.67 X_{11} $r_p = .215$ ($p = .3621$)	+ 7.52 X_{12} $r_p = .532$ ($p = .0157$)	+ 0.06 X_9 $r_p = .164$ ($p = .4885$)	28.33%	1.82

[a] After correction for serial correlation using factor of 0.40.

[b] Independent multiple regression yields:

$$X_{1\text{-}3} = 53.49 + 5.81\ X_{11} + 0.14\ X_9 + 2.96\ X_{12} + 3.89\ X_{15},\ cR^2 = 93.53\%,$$

$r_p = .853$ ($p = .0000$); $r_p = .632$ ($p = .0028$); $r_p = .688$ ($p = .0008$); $r_p = .556$ ($p = .0109$); DW = 1.83 (NS)

[c] After correction for serial correlation using factor of 0.60; not fully corrected.

[d] After correction for serial correlation using factor of 0.60; independent multiple regression yields after correction using factor of 0.40:

$$X_{1\text{-}5} = 41.45 - 7.95\ X_{13} + 4.89\ X_{12} - 0.35\ X_6 + 3.27\ X_{11},\ cR^2 = 85.19\%,$$

$r_p = -.820$ ($p = .0000$); $r_p = .661$ ($p = .0021$); $r_p = -.670$ ($p = .0017$); $r_p = .509$ ($p = .0261$); DW = 1.88 (NS)

Table 4. Replication of Regression Equation 2—Individual Weekly Output (X_2) upon Independent Variables from Table 2

X_2 for Operator(s)	Intercept	Net Hours Per Week	Managerial Discipline	Economic Depression	Scheduled Rest Time	Small Group Incentive	Variance Explan.	Durbin-Watson Coef.
1A+1[a]	−661.18	+ 72.68 X_6	+516.73 X_{11}	+320.25 X_{12}	+ 10.15 X_9	+103.24 X_{15}	91.59%	1.99
		$r_p = .963$	$r_p = .884$	$r_p = .761$	$r_p = .695$	$r_p = .220$		
		(p = .0000)	(p = .0000)	(p = .0002)	(p = .0014)	(p = .3814)		
2A+2[b]	−1173.33	+ 77.46 X_6	+635.97 X_{11}	+361.84 X_{12}	+ 9.98 X_9	+120.08 X_{15}	93.40%	1.97
		$r_p = .963$	$r_p = .942$	$r_p = .829$	$r_p = .676$	$r_p = .265$		
		(p = .0000)	(p = .0000)	(p = .0000)	(p = .0021)	(p = .2880)		
3	−641.62	+ 66.26 X_6	+277.74 X_{11}	+128.22 X_{12}	+ 7.65 X_9	+192.76 X_{15}	95.44%	2.00
		$r_p = .975$	$r_p = .891$	$r_p = .627$	$r_p = .717$	$r_p = .561$		
		(p = .0000)	(p = .0000)	(p = .0053)	(p = .0008)	(p = .0154)		
4[c]	−284.29	+ 69.22 X_6	+310.10 X_{11}	+240.27 X_{12}	+ 9.10 X_9	+ 5.42 X_{15}	90.28%	1.44
		$r_p = .962$	$r_p = .688$	$r_p = .598$	$r_p = .654$	$r_p = .012$		
5[d]	−19.70	+ 53.09 X_6	+109.48 X_{11}	+257.13 X_{12}	+ 2.57 X_9	+ 51.99 X_{15}	70.57%	1.96
		$r_p = .874$	$r_p = .218$	$r_p = .470$	$r_p = .161$	$r_p = .075$		
		(p = .0000)	(p = .3848)	(p = .0490)	(p = .5239)	(p = .7686)		

[a] After correction for serial correlation using factor of 0.40; independent multiple regression yields after correction using factor of 0.40:

$$X_{2-(1A+1)} = -667.78 + 72.72 X_6 + 520.79 X_{11} + 10.36 X_9 + 320.45 X_{12},\ cR^2 = 91.69\%,$$
$r_p = .962$, $r_p = .881$, $r_p = .695$, $r_p = .753$, DW = 1.91 (NS)
(p = .0000) (p = .0000) (p = .0010) (p = .0002)

[b] After correction for serial correlation using factor of 0.10; independent multiple regression yields:

$$X_{2-(2A+2)} = -1380.15 + 78.36 X_6 + 652.06 X_{11} + 10.93 X_9 + 365.12 X_{12},\ cR^2 = 94.26\%,$$
$r_p = .961$, $r_p = .951$, $r_p = .751$, $r_p = .833$, DW = 1.82 (NS)
(p = .0000) (p = .0000) (p = .0001) (p = .0000)

[c] After correction for serial correlation using factor of 0.60; not fully corrected. Independent multiple regression yields after incomplete correction using factor of 0.60:

$$X_{2-4} = -284.06 + 69.22 X_6 + 309.84 X_{11} + 9.09 X_9 + 240.20 X_{12},\ cR^2 = 90.85\%,\ DW = 1.44$$
$r_p = .962$, $r_p = .688$, $r_p = .654$, $r_p = .598$

[d] After correction for serial correlation using factor of 0.60; independent multiple regression yields after correction using factor of 0.40:

$$X_{2-5} = 124.38 + 48.30 X_6 - 337.95 X_{18} + 190.72 X_{12} + 174.28 X_{11},\ cR^2 = 89.13\%,$$
$r_p = .944$, $r_p = -.813$, $r_p = .620$, $r_p = .586$, DW = 1.86 (NS)
(p = .0000) (p = .0000) (p = .0046) (p = .0083)

similar to the regression results of model 1 for individual hourly output. In addition, net hours worked per week are positively associated with output for all five workers, but the model 2 variable of small group incentive payment has a significant effect only for worker 3 (see fns. to Table 4). The best individual equations show variance explanation of weekly output which ranges from 95% for operator 3 to 89% for operator 5.

Model 3 regressions of repair time (poor quality output) upon experimental variables are presented in Table 5 for individual workers. Although the group regression equation is not fully supported by any individual results, the main variable, the provision of defective raw materials, is significant and the most important variable for each individual. More frequent scheduled rest stops seem detrimental to production quality for three workers, while the economic depression and fewer weeks per experimental period seem to have contributed to poor quality for two workers. Independent stepwise multiple regression for individuals shows a negative influence of voluntary rest time upon work quality for operators 1A plus 1, negative influence of scheduled rest time for operator 3 (in place of number of scheduled rest stops), and a beneficial influence of managerial discipline upon work quality for operator 5. Best equations show vari-

Table 5. Replication of Regression Equation 3—Individual Repair Time (X_3) upon Independent Variables from Table 2

X_3 for Opera-tor(s)	Inter-cept	Defective Raw Mtls.	Scheduled Rest Stops	Economic Depression	Weeks per Period	Variance Explan.	Durbin-Watson Coef.
1A+1[a]	35.00	+ 28.85 X_{13} $r_p = .892$ ($p = .0000$)	+ 2.00 X_8 $r_p = .400$ ($p = .1563$)	+ 3.42 X_{12} $r_p = .316$ ($p = .2716$)	− 0.36 X_7 $r_p = -.452$ ($p = .1046$)	74.74%	2.06
2A + 2A[b]	21.02	+ 17.46 X_{13} $r_p = .738$ ($p = .0017$)	+ 2.72 X_8 $r_p = .544$ ($p = .0359$)	+ 4.47 X_{12} $r_p = .370$ ($p = .1750$)	− 0.30 X_7 $r_p = -.424$ ($p = .1157$)	53.38%	2.05
3[c]	21.34	+ 19.98 X_{13} $r_p = .957$ ($p = .0000$)	+ 1.85 X_8 $r_p = .695$ ($p = .0058$)	+ 12.76 X_{12} $r_p = .953$ ($p = .0000$)	− 0.05 X_7 $r_p = -.137$ ($p = .6414$)	96.16%	2.09
4[d]	6.44	+ 14.02 X_{13} $r_p = .725$ ($p = .0034$)	+ 1.07 X_8 $r_p = .293$ ($p = .3096$)	+ 10.21 X_{12} $r_p = .664$ ($p = .0096$)	− 0.25 X_7 $r_p = -.481$ ($p = .0813$)	52.44%	1.92
5[e]	9.18	+ 46.57 X_{13} $r_p = .962$ ($p = .0000$)	+ 5.58 X_8 $r_p = .865$ ($p = .0001$)	+ 4.85 X_{12} $r_p = .348$ ($p = .2231$)	− 0.50 X_7 $r_p = -.774$ ($p = .0012$)	92.08%	1.92

[a] After correction for serial correlation using factor of —0.25; independent multiple regression yields:

$$X_{3-(1A+1)} = 37.00 + 28.36\,X_{13} - 0.77\,X_{10-1} - 0.33\,X_7 \quad cR^2 = 73.54\%,$$
$r_p = .868$ ($p = .0000$); $r_p = .579$ ($p = .0187$); $r_p = -.503$ ($p = .0469$); DW = 1.99 (NS)

[b] Independent multiple regression yields:

$$X_{3-(2A+2)} = 22.41 + 15.28\,X_{13} + 3.32\,X_8, \quad cR^2 = 56.20\%, \quad DW = 2.23 \text{ (NS)}$$
$r_p = .732$ ($p = .0013$); $r_p = .583$ ($p = .0177$)

[c] After correction for serial correlation using factor of —.80; independent multiple regression yields:

$$X_{3-3} = 13.90 + 20.35\,X_{13} + 13.08\,X_{12} - 0.28\,X_7 + 0.21\,X_9, \quad cR^2 = 91.67\%,$$
$r_p = .937$ ($p = .0000$); $r_p = .915$ ($p = .0000$); $r_p = -.683$ ($p = .0050$); $r_p = .596$ ($p = .0191$); DW = 2.04 (NS)

[d] After correction for serial correlation using factor of 0.25; independent multiple regression yields:

$$X_{3-4} = 8.51 + 12.68\,X_{13} + 8.14\,X_{12}, \quad cR^2 = 58.89\%, \quad DW = 2.34 \text{ (NS)}$$
$r_p = .689$ ($p = .0022$); $r_p = .662$ ($p = .0038$)

[e] After correction for serial correlation using factor of 0.40; independent multiple regression yields after correction using factor of —0.40:

$$X_{3-5} = 34.37 + 49.09\,X_{13} + 4.68\,X_8 - 11.74\,X_{11} - 0.48\,X_7 + 8.73\,X_{12}, \quad cR^2 = 97.76\%$$
$r_p = .990$ ($p = .0000$); $r_p = .877$ ($p = .0001$); $r_p = -.887$ ($p = .0001$); $r_p = -.772$ ($p = .0020$); $r_p = .865$ ($p = .0001$); DW = 2.17 (NS)

ance explanations for repair time ranging from 98% for operator 5 down to 56% for operators 2A plus 2.

In general, the group output models (Table 2) are the best models also for individuals (Tables 3, 4, and 5), or are within a few percentage points of amended models in variance explanation. Median variance explanations for individuals using both the group and the best models are 87% for hourly output, 92% for weekly output, and 75% for repair time—somewhat lower than variance explanations for the group of 94%, 96%, and 93%. For quantity of output, the managerial intervention in replacing two workers and the advent of the economic depression were important to all individuals as well as to the group. Both factors may be viewed as exerting certain pressures upon the workers. For quality of output, quality of raw materials was of primary importance to all individuals and to the group.

These results differ starkly from most earlier descriptions of the findings of the Hawthorne experiments. The following section interprets the present results in juxtaposition to previous interpretations.

DISCUSSION

Multiple regression analyses over 23 periods of the first relay experiment at Hawthorne show that three variables—managerial discipline, the economic adversity of the depression, and time set aside for rest—explain most of the variance in quantity of output for the group and generally for individual workers. Two workers who exhibited undue independence from management (but, as shown in Appendix 2, did not have the lowest average production rates in the group) were replaced by two more agreeable workers. This exercise of managerial discipline seems to have been the major factor in increased rates of output for the now altered group, including increased production by the three individuals remaining. It may be speculated that improvement resulted from the positive example of the two new workers, as well as from the aversive effects of management's disposal of two of the original workers. Clear support is given to the suggestion by Carey (1967) that this intervention was a key part of the first relay experiment. As pointed out by Argyle (1953: 100), the Hawthorne researchers had provided "no quantitative evidence for the conclusion for which this experiment is famous—that the increase of output was due to a changed relation with supervision." Quantitative evaluation now does provide such evidence. However it is not "release from oppressive supervision," as suggested by Landsberger (1958:53), but its reassertion that explains higher rates of production.

Regarding the second independent variable resulting from the present analyses, the Hawthorne researchers as well as Argyle (1953) and Landsberger (1958) recognized that the economic depression beginning October 24, 1929, might have been a disturbing factor in the experiments. Yet they did not appear to suspect its positive influence on production. The increased importance of jobs and the real danger of losing them, because of the depression, may explain its positive contribution to output quantity for the group and for all individuals.

The third dimension resulting from these analyses is worker fatigue. In his review of the Hawthorne studies and of the first several decades of criticism and application, Landsberger (1958) agreed with the conclusion of the Hawthorne researchers that reduction of fatigue did not play much of a role in the first relay experiment. This seemed indicated to them by examination of individual work rates over time and by the findings from the mica splitting experiment. However, Argyle (1953) and Carey (1967) reviewed the same evidence and came to opposite conclusions. Their interpretations are supported by the present analyses, suggesting that physical or mental fatigue reduction through rest pauses also contributed to higher output rates for the group and for four of the five workers. In the case of worker 5, fatigue reduction by working fewer hours in the week appears to have increased the rate of hourly output. Crude analysis of data from the mica splitting experiment further suggests that reduction of fatigue is beneficial for production (fn. 9).

An additional variable found related to production is the use of an incentive pay system based upon the output of the small group rather than upon that of the department. Again, contrary to the views of the Hawthorne researchers and of most subsequent interpreters, some empirical evidence is provided here supporting the positive influence of small group incentives upon weekly output rates in the first relay experiment and upon hourly output rates in the second relay experiment (fn. 9). However, the effects of incentives in the first relay experiment are minor relative to the three factors discussed above, and thus will not be considered further.

In this statistical analysis of the first relay experiment, only the relatively small but consistent effect of rest pauses upon production quantity provides support for the contention that economic benefits result from humanitarian activity. The lack of substantial unexplained variance in any of the final models for output quantity indicates that the unmeasured supervisory and social interaction variables were not very important economically. As Carey (1967) suggested in his incisive but nonquantitative critique, reevaluation of the experiments does not support the conclusions of the Hawthorne investigators.[12] Still, there remains Carey's (1967:403) question of how it was possible for "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." One explanation for this enthusiastic embrace of something scientifically unproved may lie in the particular emphasis of the Hawthorne conclusions, another in the nonsubstantive nature of most criticism of them. Conclusions of the Hawthorne studies seem to have been congenial to persons who were in agreement with the prevailing economic system, but were prepared to proceed from simple materialistic notions about work motivation on to more complex social theories, which could be seen as more useful, humane, and democratic. Authors who appear to have interpreted the Hawthorne studies in this way include those in the volume edited by Cass and Zimmer (1975), and DeNood (1941), Friedman (1946), Homans (1941; 1949; 1950), Landsberger (1958), Miller and Form (1951), Nieder (1975), Sanford (1973), Shepard (1971), and Vroom (1964).

Most criticism in early years was ideological rather than substantive, in part directed, as noted by Landsberger (1958), against the ideology of Mayo (1919; 1933; 1945; 1947) and Whitehead (1936), and not particularly concerned with what the Hawthorne studies themselves had to say. This criticism did not treat seriously the main body of work by Roethlisberger and Dickson (1939), supplemented by Whitehead (1938). Examples of such rather misdirected interpretations include the writings of Bell (1947), Gilson (1940), Lynd (1937), Bendix and Fisher (1949), and Schneider (1950). But complaints by Sheppard (1949; 1950) and Hampden-Turner (1970) of reactionary tendencies at Hawthorne have been given some plausibility by Homans's (1941) and Wilensky and Wilensky's (1951) observations that union activities failed at Western Electric. Other social scientists have been diverted by the Hawthorne effect, described by Roethlisberger (1941:14): ". . . If a human being is being experimented upon, he is likely to know it. Therefore, his attitudes toward the experiment and toward the experimenters become very important factors in determining his responses to the situation" (cf. also Dickson and Roethlisberger, 1966, and Bishop and Hall, 1971). This concept of influence upon an experiment through the experiment itself was found either erroneous or misleading by Cook and Campbell (1976), Katz and Kahn (1966), Parsons (1974), and Rubeck (1975). Sommer's (1968) conclusion, that the "errors" called placebo or Hawthorne effect need themselves to be evaluated and understood, is most pertinent.

Perhaps discouraged by the inaccessibility of numerical data from the experiments (although outputs were graphed and most independent variables described by Whitehead, 1938), not one of the numerous commentators has attempted a quantitative interpretation of changes in output rates for nearly 50 years.[13] Indeed, except for Hare (1967) and Parsons (1974), most of the interpreters of the first relay experiment appear not to have recognized that there were more than 13 experimental periods, even though 15 periods had been described by Pennock (1930) and a total of 24 by Whitehead (1938; cf. fn. 6).

In the social sciences—particularly where complex beliefs and processes are

[12] Acker and Van Houten (1974:156) similarly concluded that "the cumulative effect of coercion, paternalistic treatment, and special rewards resulted in a rise in productivity." Others that criticized the early evaluations of the first relay experiment are Argyle (1953), Blum and Naylor (1968), Farris (1969), Locke (1976), Moore (1947), Sykes (1965), and Viteles (1953). Acker and Van Houten further suggested that results for the female workers in the first relay experiment might not be applicable to male workers.

[13] Parsons (1974) offered a behavioral theory to replace the social interaction theory of the Hawthorne investigators, but this theory also is rendered implausible by the present analyses. Still, quantitative testing of the feedback mechanism suggested by Parsons should be possible using the continuous production record available from the first relay experiment.

involved—there seems to be no substitute for quantitative analysis. Whether there are only few data available or, as in the present case, where there exists a massive body of data and description, quantitative analysis enables the scientist to separate fact from fiction. Much of the information from the Hawthorne experiments remains to be tapped and interpreted with this aim.[14]

CONCLUSION

This first statistical interpretation of the major Hawthorne experiment leads to conclusions different from those heretofore drawn. Most of the variance in production rates during the first relay experiment could be explained by measured variables. To assume that output changes resulted from unmeasured changes in the human relations of workers therefore seems injudicious, even though it was the assumption of the Hawthorne researchers and has been accepted and built upon by many social scientists over the past several decades.

The Hawthorne experiments, most of which involved small groups of workers, are exceptional in the accumulation of information over extensive periods of time under actual working conditions. The experiments drew attention to small group processes, and the studies' conclusions led to widespread acceptance of human relations as a primary factor in worker performance. Following dissemination of the findings, previously accepted and conceptually simpler mechanisms such as those of scientific management (Taylor, 1911) tended to be given less emphasis as determinants of work performance. These include the possible benefits of fatigue reduction, use of economic incentives, the exercise of discipline, and other aspects of managerial control. But it is precisely such factors to which we are directed by empirical analyses of the Hawthorne data. In particular, the discharge and replacement of two somewhat insubordinate workers were followed by higher group and individual production rates in the first relay experiment. Fairly strong evidence has been provided in recent years showing that proclivity to exert close managerial control can benefit the economic performance of individual managers (Miner, 1965), of organizations (Kock, 1965), and of whole societies (Franke, 1973; 1974; 1977). If the empirical results from the Hawthorne experiments and from these more recent studies contain some general applicability to economic organizations, then more of our attention as social scientists might well be directed to managerial characteristics and processes and somewhat less to the human relations of workers. Quantitative analyses of the data from Hawthorne, as well as empirical studies of work groups in the decades subsequent (cf. Stogdill, 1974), unfortunately do not support a contention that improvements in human relations lead to improved economic performance. On the other hand, such activities as participative management, industrial democracy, and sensitivity or consideration training may have benefits transcending the criteria considered here.

The analytical procedures employed in the present study suggest feasibility of examining closely the building blocks of our disciplines, especially when quantitative information is available. This has long been done in the physical sciences, where development routinely includes the process of critical scientific review, secondary analysis, and replication of important studies. There appears great need as well as opportunity for such activities in the social sciences.

[14] A "Guide to Hawthorne Records," which provides entry to the UWM and WPI microfilm files and to the comprehensive index of Mallach and Smith (1977), may be obtained from Franke. All data used in the present analysis are provided in Appendices 1 and 2. We wish to encourage reappraisal of our calculations, as well as further investigation of these historically important experiments.

APPENDIX 1

Group Data from First Relay Experiment [a]

Period Number (Dates Included)	(1) Hourly Output	(2) Weekly Output	(3) Repair Time	(4) Hours Per Day	(5) Days Per Week	(6) Net Hours Per Week	(7) Weeks Per Period	(8) Scheduled Rest Stops	(9) Scheduled Rest Time	(10) Voluntary Rest Time
1 (4/25–5/10/27)	49.7	2385.60	—	8.75	5.5	48.00	2	0	0	—
2 (5/10–6/11/27)	49.1	2356.80	—	8.75	5.5	48.00	5	0	0	10.5
3 (6/13–8/6/27)	51.0	2448.00	14.9	8.75	5.5	48.00	8	0	0	13.7
4 (8/8–9/10/27)	52.1	2452.87	18.5	8.75	5.5	47.08	5	2	10	9.0
5 (9/12–10/8/27)	55.1	2543.97	26.4	8.75	5.5	46.17	4	2	20	9.5
6 (10/10–11/5/27)	55.6	2515.90	31.7	8.75	5.5	45.25	4	6	30	0.5
7 (11/7/27–1/21/28)	55.9	2552.95	18.8	8.75	5.5	45.67	11	2	25	8.4
8 (1/23–3/10/28)	61.9	2692.22	23.2	8.25	5.5	43.17	7	2	25	2.8
9 (3/12–4/7/28)	63.9	2598.81	17.2	7.75	5.5	40.67	4	2	25	2.3
10 (4/9–6/30/28)	61.8	2822.41	15.8	8.75	5.5	45.67	12	2	25	5.5
11 (7/2–9/1/28)	62.8	2616.88	19.4	8.75	5.0	41.67	9	2	25	6.4
12 (9/3–11/24/28)	60.7	2913.60	13.4	8.75	5.5	48.00	12	0	0	14.3
13 (11/26/28–6/29/29)	66.5	3039.06	14.4	8.75	5.5	45.67	31	2	25	7.0
14 (7/1–8/31/29)	63.3	2637.71	48.5	8.75	5.0	41.67	9	2	25	6.9
15 (9/2/29–4/5/30)	66.2	3023.35	40.2	8.75	5.5	45.67	31	2	25	5.2
16 (4/7–5/3/30)	69.7	3183.20	30.5	8.75	5.5	45.67	4	2	25	4.6
17 (5/5–10/25/30)	69.2	2624.06	22.0	8.00	5.0	37.92	25	2	25	5.4
18 (10/29/30–2/7/31)	69.6	2406.79	22.0	8.00	4.5	34.58	15	2	25	7.1
19 (2/9–5/23/31)	69.3	2396.39	26.9	8.00	4.5	34.58	15	2	25	7.0
20 (5/25–11/14/31)	68.6	2601.31	24.2	8.00	5.0	37.92	25	2	25	6.0
21 (11/16–12/5/31)	69.6	2110.97	—	8.00	4.0	30.33	3	2	25	5.2
22 (12/7/31–2/6/32)	71.7	2718.86	—	8.00	5.0	37.92	9	2	25	4.6
23 (2/8–2/27/32)	71.5	2168.60	—	8.00	4.0	30.33	3	2	25	—

[a] For specific microfilm sources, contact first author (cf. fn. 14). Values of categoric variables: X_{11} over periods 1–7: 0, periods 8–23: 1; X_{12} over periods 1–14: 0, period 15: 0.74, periods 16–23: 1; X_{13} over periods 1–13 and 16–23: 0, periods 14–15: 1; X_{14} over periods 1–13 and 18–23: 0, period 14: 0.44, periods 15–16: 1, period 17: 0.04; X_{15} over periods 1–2: 0, periods 3–23: 1.

APPENDIX 2

Individual Data from First Relay Experiment [a]

	(1) Hourly Output, units/hour					(3) Repair Time, min./day					(10) Voluntary Rest Time, min./day				
Operator:	1A+1	2A+2	3	4	5	1A+1	2A+2	3	4	5	1A+1	2A+2	3	4	5
Period															
1	50.5	49.7	49.7	49.7	48.3	—	—	—	—	—	—	—	—	—	—
2	47.8	48.0	49.5	51.1	48.9	—	—	—	—	—	10.0	12.6	12.0	12.7	5.3
3	48.4	50.4	53.6	52.2	50.5	19.0	13.5	12.3	10.6	19.0	13.7	17.0	13.0	16.5	8.3
4	51.5	50.7	53.6	53.6	50.8	19.7	16.3	13.4	9.4	32.8	8.8	4.8	11.0	13.9	6.5
5	54.1	55.4	56.9	56.1	52.9	34.6	37.4	17.4	11.4	31.5	9.8	9.3	9.3	12.8	6.5
6	55.2	54.7	56.8	56.8	54.6	38.0	36.3	21.2	12.5	50.3	0.3	1.1	0.2	0.5	0.6
7	54.0	53.9	58.9	58.2	54.2	23.5	23.0	14.2	6.2	27.0	11.4	6.1	10.3	9.0	5.1
8	62.8	64.5	62.2	63.1	56.8	39.0	25.2	19.0	6.3	26.4	0.5	5.3	1.3	4.5	2.2
9	65.5	68.0	63.0	63.5	59.5	30.5	16.8	16.3	6.0	16.3	—	4.6	0.5	5.0	1.6
10	63.9	64.9	62.1	62.8	55.2	24.1	20.0	12.5	3.5	18.7	0.2	7.6	3.6	11.2	4.9
11	65.6	66.4	63.9	62.9	55.0	33.6	27.6	13.1	9.6	13.2	—	8.9	4.3	13.8	4.8
12	62.5	63.9	59.7	61.3	56.1	23.7	22.9	11.0	3.8	5.6	15.0	20.6	11.2	15.8	8.9
13	67.4	71.9	64.3	69.1	59 7	22.3	19.0	13.8	7.8	8.9	5.0	10.0	5.4	9.2	5.6
14	64.5	70.6	62.0	69.4	50.0	59.9	42.7	40.7	30.0	69.0	6.1	8.4	7.1	8.0	5.1
15	68.8	73.9	64.0	71.3	52.7	48.8	36.7	36.8	18.4	60.2	5.1	6.0	5.1	5.2	4.5
16	72.7	76.8	67.5	73.0	59.6	39.5	31.7	31.7	19.2	30.5	3.9	4.4	6.1	4.8	4.2
17	69.6	75.9	65.6	73.7	61.5	28.0	20.3	26.1	14.6	21.1	4.6	5.0	6.3	4.9	6.4
18	69.1	76.4	64.9	72.8	64.9	21.0	25.0	26.0	13.7	24.5	6.3	5.7	9.2	6.2	8.0
19	72.9	74.8	64.1	69.6	65.3	32.8	30.0	30.6	19.6	21.3	6.1	5.9	8.3	6.9	7.6
20	72.5	76.3	63.4	66.5	64.0	28.7	25.0	26.0	22.7	18.6	5.0	5.4	6.8	6.6	6.4
21	70.2	72.5	67.0	68.9	67.7	—	—	—	—	—	3.7	4.6	6.2	5.5	5.8
22	75.0	76.4	68.6	73.1	65.8	—	—	—	—	—	3.6	3.6	5.0	5.1	5.6
23	77.1	78.0	66.9	73.1	67.1	—	—	—	—	—	—	—	—	—	—

[a] For specific microfilm sources, contact first author (cf. fn. 14). Individual weekly output rates can be calculated by multiplying individual values of variable (1), above, by variable (6) from Appendix 1.

REFERENCES

Acker, Joan and Donald R. Van Houten
1974 "Differential recruitment and control: the sex structuring of organizations." Administrative Science Quarterly 19:152-63,

Argyle, Michael
1953 "The relay assembly test room in retrospect." Occupational Psychology 27:98-103.

Bass, Bernard M.
1960 Leadership, Psychology, and Organizational Behavior. New York: Harper.

Bass, Bernard M. and Gerald V. Barrett
1972 Man, Work, and Organizations. Boston: Allyn and Bacon.

Beer, Michael
1976 "The technology of organizational development." Pp. 937–93 in M. D. Dunnette (ed.), Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally.

Bell, Daniel
1947 "Adjusting men to machines." Commentary 3:79–88.

Bendix, Reinhard and Lloyd H. Fisher
1949 "The perspectives of Elton Mayo." Review of Economics and Statistics 31:312–9.

Bishop, Ronald C. and James W. Hill
1971 "Effects of job enlargement and job change on contiguous but nonmanipulated jobs as a function of workers' status." Journal of Applied Psychology 55:175–81.

Blalock, Hubert M., Jr.
1961 Causal Inferences in Nonexperimental Research. Chapel Hill: University of North Carolina Press.

Blum, Milton L. and James C. Naylor
1968 Industrial Psychology: Its Theoretical and Social Foundations. New York: Harper and Row.

Bradford, L. P., J. Gibb and K. Benne (eds.)
1964 T-group Theory and Laboratory Method. New York: Wiley.

Campbell, Donald T. and Julian C. Stanley
1963 "Experimental and quasi-experimental designs for research on teaching." Pp. 171–246 in N. L. Gage (ed.), Handbook of Research on Teaching. Chicago: Rand McNally.

Carey, Alex
1967 "The Hawthorne studies: a radical criticism." American Sociological Review 32:403–16.

Cass, Eugene Louis and Frederick G. Zimmer (eds.)
1975 Man and Work in Society. New York: Van Nostrand Reinhold.

Cherrington, David J., H. Joseph Reitz, and William E. Scott, Jr.
1971 "Effects of contingent and non-contingent reward on the relationship between satisfaction and task performance." Journal of Applied Psychology 55:531–6.

Cook, Thomas D. and Donald T. Campbell
1976 "The design and conduct of quasi-experiments and true experiments in field settings." Pp. 223–326 in M. D. Dunnette (ed.), Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally

DeNood, Neal B.
1941 "Book review: management and the worker." American Sociological Review 6:304–5.

Dickson, William J. and F. J. Roethlisberger
1966 Counseling in an Organization. Boston: Harvard University.

Dogan, M. and S. Rokkan (eds.)
1969 Quantitative Ecological Analysis in the Social Sciences. Cambridge, Ma.: MIT Press.

Durbin, J. and G.S. Watson
1950 "Testing for serial correlation in least squares regression, I."Biometrika 37:409–28.
1951 "Testing for serial correlation in least squares regression, II." Biometrika 38:159–78.

Elliott, J. Walter
1973 Economic Analysis for Management Decisions. Homewood: Irwin.

Farris, George F.
1969 "The drunkard's search in behavioral science." Personnel Administration 32:10–8.

Fleishman, E. A., E. F. Harris, and H. E. Burtt
1955 Leadership and Supervision in Industry. Colombus: Bureau of Educational Research, Ohio State University.

Franke, Richard Herbert
1973 "Critical factors in the post-war economic growth of nations: review of empirical studies and implications for participative organization." Pp. 107–19 in E. Pusić (ed.), Participation and Self Management, Vol. 5. Zagreb: University of Zagreb.
1974 An Empirical Appraisal of the Achievement Motivation Model Applied to Nations. Ph.D. dissertation, Graduate School of Management, University of Rochester.
1977 "Culture and industrial development." Paper presented at the International Conference on Social Change and Organizational Development, Inter-University Centre of Post-Graduate Studies, Dubrovnik.

Friedman, Georges
[1946] 1955 Problèmes Humains du Machinisme Industriel. Paris: Librairie Gallimard. Trans. and ed. by Harold L. Sheppard as Industrial Society: the Emergence of the Human Problems of Automation. Glencoe: Free Press.

Galtung, Johan
1967 Theory and Methods of Social Research. New York: Columbia University Press.

Gilson, Mary B.
1940 "Book review: management and the worker." American Journal of Sociology 46:98–101.

Golembiewski, Robert T. and Arthur Blumberg
1970 Sensitivity Training and the Laboratory Approach. Itasca: Peacock.

Hampden-Turner, Charles
1970 Radical Man. Cambridge, Ma.: Schenkman.

Hare, A. Paul
1967 "Small group development in the relay assembly testroom." Sociological Inquiry 37:169–82.
Homans, George Caspar
1941 "The Western Electric researches." Pp. 56–99 in National Research Council Committee on Work in Industry, Fatigue of Workers: Its Relation to Industrial Production. New York: Reinhold.
1949 "Some corrections." Review of Economics and Statistics 31:319–21.
1959 The Human Group. New York: Harcourt Brace.
Johnston, J.
1963 Econometric Methods. New York: McGraw-Hill.
Kahn, Robert L.
1975 "In search of the Hawthorne effect." Pp. 49–63 in Eugene Louis Cass and Frederick G. Zimmer (eds.), Man and Work in Society. New York: Van Nostrand Reinhold.
Katz, Daniel and Robert L. Kahn
1966 The Social Psychology of Organizations. New York: Wiley.
Kock, Sven E.
1965 Företagsledning och motivation. (Management and Motivation, with English summary.) Helsinki: Affärsekonomisk Förlagsförening.
Landsberger, Henry A.
1958 Hawthorne Revisited. Ithaca: Cornell University.
Likert, Renis
1967 The Human Organization. New York: McGraw-Hill.
Locke, Edwin A.
1976 "The nature and causes of job satisfaction." Pp. 1297–349 in Marvin D. Dunnette (ed.), Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally.
Lynd, Robert S.
1937 "Book review: leadership in a free society." Political Science Quarterly 52:590–2.
Mallach, Stanley and Steven Smith
1977 Records of the Industrial Relations Experiment Carried Out by the Western Electric Company at the Hawthorne Works, Hawthorne, Illinois. Index, University of Wisconsin-Milwaukee Library, Milwaukee.
Marrow, Alfred J.
1975 "Management by participation." Pp. 33–48 in Eugene Louis Cass and Frederick G. Zimmer (eds.), Man and Work in Society. New York: Van Nostrand Reinhold.
Mayo, Elton
1919 Democracy and Freedom. Melbourne: Macmillan.
1933 The Human Problems of an Industrial Civilization. New York: Viking.
1945 The Social Problems of an Industrial Civilization. Boston: Harvard University.
1947 The Political Problems of an Industrial Civilization. Boston: Harvard University.
Miller, Delbert C. and William H. Form
1951 Industrial Society. New York: Harper.
Miner, John B.
1965 Studies in Management Education. New York: Springer.
Moore, Wilbert E.
1947 "Current issues in industrial sociology." American Sociological Review 12:651–7.
Nieder, Peter
1975 "Zum zusammenhang zwischen führungsverhalten, produktionsniveau und zufriedenheit." Gruppendynamik 2:127–39.
Parsons, H. M.
1974 "What happened at Hawthorne?" Science 183:922–32.
Pennock, G. A.
1930 "Industrial research at Hawthorne: an experimental investigation of rest periods, working conditions and other influences." Personnel Journal 8:296–313.
Price, James L.
1968 Organizational Effectiveness: An Inventory of Propositions. Homewood: Irwin.
Pusić, E. (ed.)
1973 Participation and Self Management. 6 Vols. Zagreb: University of Zagreb.
Robinson, W. S.
1959 Ecological correlations and the behavior of individuals. American Sociological Review 15:352–7.
Roethlisberger, F. J.
1941 Management and Morale. Cambridge, Ma.: Harvard University Press.
Roethlisberger, F. J. and William J. Dickson
1939 Management and the Worker. Cambridge, Ma.: Harvard University Press.
Rubeck, Patricia A.
1975 "Hawthorne concept–does it affect reading progress?" The Reading Teacher 28:375–9.
Sales, S. M.
1966 "Supervisory style and productivity: review and theory." Personnel Psychology 19:275–86.
Sanford, Aubrey C.
1973 Human Relations: Theory and Practice. Columbus: Merrill.
Schneider, Louis
1950 "An industrial sociology–for what ends?" Antioch Review 10:407–17.
Shepard, Jon M.
1971 "On Alex Carey's radical criticism of the Hawthorne studies." Academy of Management Journal 14:23–32.
Sheppard, Harold L.
1949 "The treatment of unionism in 'managerial sociology'." American Sociological Review 14:310–3.
1950 "The social and historical philosophy of Elton Mayo." Antioch Review 10:396–405.
Sommer, Robert
1968 "Hawthorne dogma." Psychological Bulletin 70:592–5.
Stogdill, Ralph M.
1974 Handbook of Leadership. Glencoe: Free Press.
Sykes, A. J. M.
1965 "Economic interest and the Hawthorne researches." Human Relations 18:253–63.

Taylor, Frederick W.
1911 The Principles of Scientific Management. New York: Harper.
Theil, H. and A. L. Nager
1961 "Testing the independence of regression disturbances." American Statistical Association Journal 56:793–806.
Thorndike, E. L.
1939 "On the fallacy of imputing the correlations found for groups to the individuals or smaller groups composing them." American Journal of Psychology 52:122–4.
Viteles, Morris S.
1953 Motivation and Morale in Industry. New York: Norton.
Vroom, Victor H.
1960 Some Personality Determinants of the Effects of Participation. Englewood Cliffs: Prentice-Hall.
1964 Work and Motivation. New York: Wiley.
Whitehead, T. N.
1936 Leadership in a Free Society. Cambridge, Ma.: Harvard University Press.
1938 The Industrial Worker. 2 Vols. Cambridge, Ma.: Harvard University Press.
Whyte, William F.
1955 Money and Motivation: An Analysis of Incentives in Industry. New York: Harper and Row.
Wilensky, Jeanne L. and Harold L. Wilensky
1951 "Personnel counseling: the Hawthorne case." American Journal of Sociology 57:265–80.

THE MYTH OF SOCIAL CLASS AND CRIMINALITY: AN EMPIRICAL ASSESSMENT OF THE EMPIRICAL EVIDENCE*

CHARLES R. TITTLE
Florida Atlantic University

WAYNE J. VILLEMEZ
University of Illinois, Chicago Circle

DOUGLAS A. SMITH
Indiana University

American Sociological Review 1978, Vol. 43 (October):643–656

Thirty-five studies examining the relationship between social class and crime/delinquency are reduced to comparable statistics using instances where the relationship was studied for specific categories of age, sex, race, place of residence, data type, or offense as units of analysis. The findings from 363 instances are summarized and patterns are identified. The overall results show only a slight negative relationship between class and criminality, with self-report studies reflecting lower associations than official statistics studies. Moreover, analysis demonstrates a clear historical decline in magnitude of association to the point where both self-report and official statistics studies done in the current decade find no class variation. This historical trend is shown to be due to changes in the findings from studies using official statistics as indicators of criminality. Alternative interpretations are discussed, but all lead to serious doubts about the adequacy of theories of deviance that contain assumptions of class differences.

Social scientists long have assumed an intimate linkage between a variety of social pathologies and injustice or inequity in the distribution of societal resources. This is a reasonable assumption because differences in social power and advantage imply differences across the whole range of life chances. But a relationship between the distribution of social resources and behavioral manifestations is more easily justified on theoretical than empirical grounds. For one thing, concentration of resources into distinguishable categories never has been measured clearly enough to permit firm conclusions about relationships. Indeed, controversy about the extent of resource concentration has pervaded the stratification literature. At one point social class was a widely accepted concept for describing such concentrations, but following a concerted attack in the late fifties and early sixties (e.g., Cutright, 1968; Glenn, 1967; Laumann, 1966; Ossowski, 1963; Nisbet, 1959; Rose,

* Address all communications to: Charles R. Tittle; Department of Sociology/Social Psychology; Florida Atlantic University; Boca Raton, FL 33431.