4-1990

The Effects of Performance Feedback on the Implementation of a Statistically-Based Quality Control Program

Gordon O. Henry
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses
Part of the Industrial and Organizational Psychology Commons

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/1040

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE EFFECTS OF PERFORMANCE FEEDBACK ON THE IMPLEMENTATION OF A STATISTICALLY-BASED QUALITY CONTROL PROGRAM

by

Gordon O. Henry

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
April 1990

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
THE EFFECTS OF PERFORMANCE FEEDBACK ON THE IMPLEMENTATION OF A STATISTICALLY-BASED QUALITY CONTROL PROGRAM

Gordon O. Henry, M.A.
Western Michigan University, 1990

Although various types of performance feedback have been shown to be effective in maintaining work-related behaviors in numerous settings, most of these behaviors have consisted of fairly simple tasks. More specifically, it has not been conclusively shown that such feedback procedures can be used to maintain the worker behaviors required in the implementation of a statistically-based quality control program. The present study attempted to show that such complex behaviors could be maintained using effective feedback procedures.

The results showed that the subjects (machine operators) performed at a high level in completing required tasks associated with a statistically-based quality control program when feedback was present. Also, the subjects performed at a slightly lower level, on the average, when feedback was not present. Although the overall quality of finished products was not a dependent variable, this measure was tracked in order to show that overall quality of products did not suffer when feedback was implemented.

It was concluded that supervisory feedback which is timely and presented in an easily understood form can be effective in maintaining fairly complex work-related behaviors like those involved in the implementation of a quality control program.
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
The effects of performance feedback on the implementation of a statistically-based quality control program

Henry, Gordon O., M.A.
Western Michigan University, 1990
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. METHODS</td>
<td>9</td>
</tr>
<tr>
<td>Subjects</td>
<td>9</td>
</tr>
<tr>
<td>Setting</td>
<td>9</td>
</tr>
<tr>
<td>Apparatus/Materials</td>
<td>10</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>11</td>
</tr>
<tr>
<td>Reliability</td>
<td>15</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>16</td>
</tr>
<tr>
<td>Procedures</td>
<td>17</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>18</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>20</td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>26</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>A. Informed Consent Form</td>
<td>32</td>
</tr>
<tr>
<td>B. Operator Sampling Plan</td>
<td>34</td>
</tr>
<tr>
<td>C. Human Subjects Institutional Review Board Approval</td>
<td>36</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>38</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Percent Measurements Within 50% Control Limits and
   Within 25% Limits................................................................. 24
# LIST OF FIGURES

1. Operator Control Chart ................................................................. 12
2. Checklist of Daily Operator Tasks Completed ................................. 14
3. Percentage of Checklist Tasks Completed Each Day for All Operators ........................................................................ 21
CHAPTER I

INTRODUCTION

Dr. W. E. Deming (1975) wrote that the poor quality control (QC) systems used by American industries and service organizations have had far-reaching economic and social ramifications. According to Deming, poor quality of goods and services have contributed to the downfall of the dollar and to the uneven balance of payments experienced in the United States. Deming maintained that the systems commonly used in the U.S. which attempted to sort out a defective product after it had been produced were not adequate. He advised that the rest of the world follow Japan's lead and adopt QC programs which concentrate on defects in the manufacturing process. Other quality experts also propose that U.S. companies forego end-of-the-line quality programs and concentrate on the quality of the process that produces the product (see Lowe & Mazzeo, 1986, for examples).

It seems that American firms have followed this advice in great earnest. Now in the United States (and around the world) there are almost as many process-oriented quality control programs as there are organizations interested in upgrading the final quality of their products (see Dumas, Cushing, & Laughlin, 1987; Lowe & Mazzeo, 1986; Redmon & Dickinson, 1987; or Vardeman & Cornell, 1987, for partial reviews). In addition, there are a great number of theoretical bases upon which most of the existing programs are predicated. Although each theory takes a slightly different approach to quality, three models will be briefly outlined which enjoy widespread use and are typical of the field in general: the Crosby model, Statistical
Process Control (SPC), whose main proponent is W. E. Deming, and the Taguchi method.

Like the other two models, the quality control program designed by Philip Crosby has as its emphasis the improvement of the manufacturing process rather than the inspection of finished goods. Crosby has developed a 14-step outline which companies should use when implementing his program (see Lowe & Mazzeo, 1986, for a more detailed description). Most of these steps concern improving attitudes pertaining to quality throughout the organization. For example, Step 1 is to get management commitment to quality improvement. Step 2 is to form quality improvement teams. The outcome is a structured approach to launching the improvement process and changing the quality-related culture of the organization.

Deming's SPC model of quality control is similar to that of Crosby's in that the emphasis is on the proper structuring of the organization so that improvements in quality can be realized (see Lowe & Mazzeo, 1986, or Redmon & Dickinson, 1987, for a more detailed analysis). Deming, too, has 14 steps in the implementation of his program. Again, these steps address needed changes in the organization prior to the beginning of the manufacturing process. For example, Step 1 explains how to create a constancy of purpose throughout the organization, Step 4 requires suppliers to provide statistical evidence of the quality of incoming materials, Step 6 describes the required employee training, and Step 7 says to provide the employees with the proper tools. Like Crosby, Deming advocates a top-down process of QC which requires a participative management style. Deming does, however, place a much greater reliance upon statistical proof that the manufacturing process will yield quality products. This proof is derived from control charts which Deming has developed. Such charts typically include the blueprint dimensions for a particular manufactured product and a range within which products will be considered acceptable. This range is bounded by
upper and lower control limits. The actual machine operators are required to complete the control charts during the manufacturing process, and, after performing some mathematical computations, the operators are able to determine if the manufacturing process is "in control" or will continue to result in high-quality products.

Dr. Genichi Taguchi utilizes even more elaborate statistics in his system of quality control than those found in Deming's SPC. In fact, Taguchi graphs and charts are often used to make more detailed analyses of data than those provided by the SPC control charts. Only three primary steps have been identified in Taguchi's program design (Gunter, 1987). Step 1 is systems design in which the concept of quality improvement is set. Step 2 is parameter design which sets the targets of the program. Step 3 is called allowance design and involves the setting of tolerances within the quality program. Again, the implementation of this program is aimed primarily at the planning stages of the manufacturing process.

All three of these models make extensive use of instructions and training programs which occur long before the actual manufacturing process begins. In most cases, management and on-line workers attend months of regular meetings and verbal presentations which describe the respective duties of each employee as they apply to the overall QC program. This reliance upon the variables which precede the actual implementation of a quality control program may lead quality practitioners to ignore critical factors which influence worker behaviors: consequences.

In fact, these models, especially Deming's SPC, have been criticized for their overreliance on antecedents of behavior to the exclusion of consequences. Redmon and Dickinson (1987) wrote that, "adding a consequence-based analysis to SPC could provide a more refined means of controlling employee behavior" (p. 63). Other authors also advocate a greater emphasis on the consequences of employee participation in a QC program. For example, Krigsman and O'Brien (1987) wrote
that employees will be made to feel pride in workmanship only if the employee has data on how well he or she is performing and management reinforces quality performance. They note that even the Japanese have begun to recognize the need to build reinforcement and feedback into the workplace because their quality programs will not survive unless they are rewarding for their participants.

Behavior analysts place a high value on the consequences of worker behaviors in establishing effective motivational conditions. Fellner and Sulzer-Azaroff (1984) stressed the importance of delivering a consequence contingent on meeting a goal for improving performance. Balcazar, Hopkins, and Suarez (1986) reviewed articles in which feedback and other consequences were presented to workers subsequent to work-related behaviors and found consistent improvements in performance. In addition, Komaki, Collins, and Penn (1982) found significant performance improvements in all departments of a processing plant only when a feedback consequence was used to bolster an antecedent condition designed to improve performance. These authors concluded that performance consequences play an important part in work motivation and that antecedents alone may not be effective in improving performance in all cases.

It is apparent that the consequences provided for a variety of work-related behaviors can be utilized to improve performance. Several of these behaviors are related to the quality of the goods produced by an organization. Two diverse courses could be followed to provide consequences for such quality-related behaviors. First, behaviors required of managers and employees in implementing quality control procedures could be followed by specified consequences. Such behaviors would include sampling outcomes according to the quality plan, keeping accurate records of sampling outcomes, and completing required graphs and charts associated with the quality program. Second, the production of high quality outputs could be directly
consequated without regard to the specific behaviors involved in following a program designed to produce quality goods. It is assumed that this type of system would increase those desirable behaviors in the manufacturing process directly related to producing quality goods. Such behaviors would include keeping machines in good repair, sharpening machine tools at appropriate times, and adjusting various settings on machines as needed.

A great deal of attention has been given to directly consequating the manufacturing of high quality finished products. Sprague, Zinn, and Kreitner (1976) describe a successful quality improvement experiment conducted by a supervisor using ideas from behavioral psychology (Luthans & Kreitner, 1975). The first-line supervisor primarily used praise as a consequence for producing quality parts. No consequences were provided for following the procedures described in the QC program, only for the ultimate production of good parts. Similarly, Allen (1987) described a study in which employees were publicly recognized at periodic meetings for producing fewer numbers of defective parts. Allen included favorable comments from customer companies who had received quality parts in his intervention. Again, no consequences were provided for following the procedures described in the quality control program. Kukla (1986) describes a very similar program in which favorable feedback from a customer company concerning the quality of certain parts was relayed to the employees involved in their production. In an interesting study in a service-oriented firm, Mundy, Passarella, and Morse (1986) used publicly displayed graphs as feedback concerning on-time deliveries for drivers of a limousine service. Once more, consequences for following the detailed behaviors outlined in a three-step implementation procedure for the quality program were ignored.

Relatively little attention has been given to techniques to motivate employees to carry out procedures that might result in the more effective implementation of QC
procedures. Bennett (1985) and Karabatsos (1985) described Elco Industries' program for rewarding employee participation in SPC training. However, no rewards were provided for actually implementing the procedures taught. Barbour (1984) described a quality program used by Simpson Industries in which employees were provided with feedback and reinforcement when they showed they had mastered the procedures involved in its implementation. But, no rewards were forthcoming if the workers actually used the procedures on the line.

Deming (1975) and others seem to assume that employees will implement the complex procedures involved in QC programs appropriately and tend to focus on the final changes in product quality. However, due to the number, complexity, and relatively abstract nature of the behaviors required of employees in implementing these programs, this assumption may be overly optimistic (see Barbour, 1984; Deming, 1975; or Lowe & Mazzeo, 1986). Levi and Mainstone (1987) wrote that employees may resist engaging in behaviors such as graphing and analyzing summary data because they seem so abstract and removed from the actual product or service that is being produced.

Given the scarcity of studies designed to maintain the behaviors involved in the implementation of QC programs, the present investigation focused on methods of managing employee performance as part of carrying out quality control procedures. This is a primary requirement since effective implementation is necessary before the procedures can be maintained long enough to change the quality of outputs.

The demands of implementation appear to be most stringent in QC programs which are statistically-based because employees are required to take various measurements from a large number of outputs, complete complex sampling charts, plot data points, perform intermediate level mathematical computations, and summarize data using statistical techniques. This type of program also requires that
these behaviors be exhibited throughout a workday. An SPC program which contains all of these implementation components will be the focus of this study.

Some simple behavior modification programs have been used as part of QC programs (see above for examples). However, few have used the most powerful technology as applied in the past few years by Organizational Behavior Managers. Frederiksen (1982) described numerous studies in which behavior modification techniques such as incentives and performance feedback were successful in maintaining a variety of employee behaviors such as those involved in production/task completion, training and development, absenteeism and tardiness, safety, and conservation. Further, he advocated the use of these techniques to address problems in other classes of work-related behaviors. Similarly, in a review, Andrasik, Heimberg, and McNamara (1981) listed nearly 100 studies in which behavior modification techniques were successful in maintaining work-related behaviors in business, industry, and government. Merwin, Thomason, and Sanford (1989) updated the work done by Andrasik et al., and concluded that organizational behavior management techniques continue to be used to impact a wide range of behaviors in the private sector.

The most powerful Organizational Behavior Management (OBM) feedback package appears to include specific skills training, well-specified goals, and performance-based feedback. With respect to QC programs, as previously mentioned, there is a much heavier emphasis upon the first two components of this package. That is, although training and goal setting are common in the implementation of quality control programs, consequences, including feedback, have not been used often.

Recent OBM research has demonstrated that performance feedback is a very powerful intervention in affecting worker behaviors. Ford (1984) compared three
feedback procedures as to their effect on improving the teaching skills of paraprofessionals in a mental retardation facility. His findings indicated that a combination of videotape and supervisory feedback resulted in the greatest and most rapid improvement in work performance. Jones, Morris, and Barnard (1986) used an instruction-and-feedback package to increase the staff completion of required forms in a psychiatric emergency room. This package consisted of individualized training and group feedback via weekly graphs. In addition, Balcazar et al. (1986) reviewed the use of performance feedback and identified general guidelines for the effective use of feedback. They recommended individual feedback, presented at least weekly, by a supervisor, in some quantitative form as the most effective feedback system. Similarly, Duncan and Bruwelheide (1986) advocate analyzing feedback procedures prior to their implementation so they may be designed to most effectively function as both reinforcers (delivered fairly immediately, by a supervisor, on an individual basis) and discriminative stimuli (presented quantitatively, in an easily understood form).

In the present study, these recommendations were used to develop a feedback system designed to aid in implementing an SPC program in a manufacturing shop. This study was also designed to evaluate specific means of motivating workers to carry out a QC program so that output quality might be improved, and so that beneficial effects might be maintained.
CHAPTER II

METHODS

Subjects

The subjects in the study consisted of three machine operators of a small metal-part processing company. All subjects were male and ranged in age from 24 to 37 years. Two of the operators selected as subjects were high school graduates, and the third subject had completed one year of college while working toward a bachelor's degree. The range in work experience was much greater. One subject had operated the type of machinery used in the study for one and one-half years. A second subject had been working on similar machinery for 6 years. The third subject had 15 years experience in machine operations. All subjects were made aware of the study and consented to the use of data for research purposes (Appendix A).

The three operators selected as subjects were chosen due to the relative uniformity of their work schedules. That is, these three operators tended to have more extended production runs and therefore had less "downtime" than the other machine operators in the firm. This was beneficial to the study in that fairly consistent, daily data were needed in order to make useful comparisons across experimental phases. Even so, as will be seen in the Results section, there were many days when data were not available from at least one subject.

Setting

The study took place in a small metal-working firm in Kalamazoo, Michigan. The company was located in a small business incubator which meant that it shared
some facilities such as conference rooms and copying machines with other small businesses existing in the same building. The company had nine full-time (including the company owner) and two part-time machine operators at the time of the study. The clerical staff consisted of two secretaries and a receptionist.

The machining operations took place in a large room having floor space of approximately 1500 square feet. The work area was kept free of obstacles and dirt, was very well lighted, and well ventilated. The noise level in the work area required that workers speak quite loudly to be heard over the running machines; however, the operators were not required to wear earplugs. All operators were required to wear protective eyewear.

**Apparatus/Materials**

The machines involved in the study consisted of two computerized metal lathes and one computerized drill press. The lathes were referred to by the subjects as LB15 and LB12, which represented the respective model numbers. The drill press was called a "mill" by the subjects. Therefore it seemed expedient to refer to the subjects according to the name of the machines they ran. This was possible because, although all subjects were trained in the operation of all three machines at the time they began working, never during the study did any subject operate more than one type of machine.

In addition to the large machines, hand-held tools were used by the subjects. Calipers and micrometers were used for taking measurements. All subjects had at least two years' experience in using these types of measurement tools.
A quality control program was established at the experimental site which adhered to the basic concepts of statistical process control. The percentage of tasks completed each day by the operators in accordance with the guidelines of this program served as the dependent variable. The subjects used two types of forms in the completion of these daily tasks. Appendix B contains an example of a sampling plan which the operators used in determining the parts on which they would make measurements of critical dimensions. Included in this plan are the batch sizes or number of parts being produced in production runs. The sampling plan also informs the operators of the specific parts (according to part number) which are to be sampled.

The second type of form used by the subjects is shown in Figure 1. This is an example of a control chart upon which are plotted the measurements of critical dimensions of metal parts taken by the operators. The sample numbers at the bottom of this chart correspond to the sample numbers at the top of the sampling plan. LCL and UCL represent the lower and upper control limits or the lower and upper ranges beyond which measurements should not be observed. If measurements were recorded beyond these limits, operators should take action such as sharpening a machine tool or adjusting the machine settings in order to bring the measurements back within the control limits. The specific action taken when measurements were beyond these limits was recorded at the bottom of the control chart.

Print dimension, at the middle left of the page, represented the exact specification for any one measurement found on the blueprint for that part. Numbers were placed in the leftmost column of the chart and represented the possible measurements for one critical dimension in inches. In the cells of the chart were placed numbers representing the actual measurements taken by the operators. Each sample consisted of measuring three consecutive parts, so the numbers 1, 2, and 3 were placed in the
Figure 1. Operator Control Chart
proper sample column and in the row containing the appropriate measurement.

In Figure 1, the measurements of 3 samples have been recorded. Notice that one numeral is circled in each sample. The circled numeral represents the median measurement of that sample. Operators were instructed to circle the median measurement in this way so that, in addition to observing if any one measurement was beyond the control limits, they could recognize trends toward such outlying measurements. By simply connecting these circled numerals, a graph line was plotted from which such trends could be inferred.

The percentage of tasks completed daily by the subjects was recorded on an experimenter checklist by the researchers. This checklist is presented in Figure 2.

Items 2 through 7 are the tasks which the operators were asked to complete on a daily basis, while items 1, 8, and 9 represent aspects of the quality control manager's performance. Thus, overall, the operators were asked to complete six tasks, represented by items 2 through 7 on the checklist, on a daily basis.

Item 2: Refers to whether or not the appropriate sampling plan (Appendix B) and the control charts (Figure 1 - one for each critical dimension to be measured) were in possession of the subject on that day.

Item 3: The subject must have tagged (using post-it stickers) the sampled parts with numbers corresponding to numbers from the sampling table.

Item 4: The operator's samples must have consisted of measurements from three consecutive parts.

Item 5: The numbers 1, 2, and 3 should have been placed in the proper columns representing each sample on the control charts.

Item 6: The median of the measurements of a critical dimension from the three sampled parts must have been circled on the control chart.
Item 7: The measurement taken by the operator of any critical dimension must have been no more than two deviations (rows on the control chart) from the same measurement on the same part taken by an experimenter who served as an independent observer. All experimenters who took part in the study underwent training in measuring techniques as described in the Reliability section below.

EXPERIMENTER CHECKLIST

Yes No N/A

______ 1. Incoming quality control conducted and recorded accurately. From supervisor’s incoming/outgoing sampling plan.

______ 2. Sampling plan in possession of operator.

______ 3. Numbered tags on most recent sample correspond to number on sampling table.

______ 4. Most recent sample consists of correct number of parts. From control sheet and operator’s bin.

______ 5. All samples recorded on control chart correctly (using 1, 2, and 3).

______ 6. Medians of all samples circled.

______ 7. Experimenter measurement of critical dimension within two increments of operator’s measurement.

Operator msmt.__________ Experimenter msmt.__________

______ 8. Outgoing quality control conducted and recorded accurately. From supervisor’s incoming/outgoing sampling plan.

______ 9. In-process quality control checklist completed and initialed before 5pm. From supervisor’s clipboard.

#Yes______ #Yes & No_______ #Yes/#Yes & No_____

Note here any parts returned or rejected due to workmanship errors:

Date__________ Part#________________________ Observer Init.__________

Figure 2. Checklist of Daily Operator Tasks Completed
Reliability

Three researchers participated in the study and all underwent reliability training prior to the start of the intervention. This training consisted of previewing the items on the experimenter checklist and agreeing on where to get the information needed to complete these items. Mastery of this portion of the training was reached when an experimenter recited where the information could be found to complete every item on the checklist. In addition, further training was required in order for the researchers to be able to complete item 7 of the checklist. This training consisted of each experimenter practicing taking measurements of both standard metal parts with known dimensions and parts selected randomly from the production floor of the machine shop with both calipers and micrometers. Mastery was achieved when a researcher's measurements were no more than one one-hundredth of an inch from the actual dimensions of a metal dowel used as an industry standard for ten consecutive trials. Also, each experimenter was required to agree (be within one one-hundredth of an inch) with another experimenter's measurements of a dimension of a selected part for five consecutive trials before training was completed.

A schedule for reliability checks was set up wherein, during periodic measurement sessions, two experimenters, independent of each other, would complete the experimental checklists. Percent agreement between experimenters was calculated using the following formula:

\[
\frac{\text{Number of Agreements}}{\text{Number of Agreements} + \text{Disagreements}} \times 100
\]

An agreement was scored when both researchers recorded the same outcome on the checklist for an item, and a disagreement was defined as a discrepancy between the experimenters on a checklist item.
Independent Variable

The independent variable in the study consisted of a feedback procedure in which the operators were provided with immediate knowledge of their performance. In this procedure, the shop foreman, who also served as the quality control manager, presented each operator with a 3 x 5 index card summarizing his performance immediately after each measurement session.

The index card contained three types of information. First, each operator was informed of the score he received pertaining to his performance that day. This score was the percentage of tasks the operator had completed correctly on the checklist that day. Next on the index card was a verbal rating corresponding to the score received. If an operator scored between 90 and 100, his verbal rating was "excellent." A score between 80 and 90 earned a "good" rating; scores between 70 and 80 were rated as "fair"; scores between 60 and 70 were "poor"; and scores from 0 to 60 were rated as "unsatisfactory." The third type of information on the index card consisted of suggestions designed for the subject which, if followed, would lead to a higher score in the future. An example of such a suggestion was, "Please remember to tag the sampled parts."

The quality control manager spoke very little to the operators concerning their daily performance at the time the cards were distributed. To insure this, and to make certain that the subjects consistently came in contact with the independent variable, a researcher sometimes accompanied the manager. In forty percent of the measurement sessions, a researcher was present with the manager when he gave the index cards to the operators. No observations were recorded in which the manager failed to give a card to a subject supposed to receive one, nor did the manager ever make verbal comments concerning the performance of the subjects during these observation periods.
Procedures

A description of the study was first presented to the operators in a meeting at which the shop foreman/quality control manager and the company owner were present. All subjects were made aware of the expense accrued when parts had to be reworked or scrapped due to quality discrepancies. In addition the owner expressed the need for a formal quality control program in order to acquire more lucrative work contracts. Also at this meeting, quality control program handbooks were distributed describing the general type of program to be instituted and its theoretical and functional bases. In addition, the specific tasks to be assigned to the operators were discussed. The quality control manager was briefed on his duties in delivering the feedback, and the meeting was adjourned.

Approximately two weeks later a "dry run" was completed in which the operators completed their assigned tasks, the manager delivered feedback, and the researchers collected the needed data. The program began with the next production run two days later.

During the program, subjects were asked to complete their control charts with measurements taken from parts according to the sampling plan for one hour each day. During the first week, parts were measured from 10 a.m. to 11 a.m. in the morning. During the second week of the project, measurements were taken from 1 p.m. until 2 p.m. in the afternoon. These measurement times were alternated each week thereafter. At the end of the hour experimenters would enter the machine shop and record on the checklist those tasks which had been completed accurately and those that had not. Since the operators took a fifteen-minute break at 11 a.m. and at 2 p.m. each day, they were not present when researchers were completing the checklists. The operators simply left their completed control charts and the sampling plans at their respective work stations when they left for break. After the checklists had been
completed by the experimenters, an index card was filled out for each operator and given to the quality control manager. The manager then presented these cards to the subjects upon their return from break.

Experimental Design

A multiple baseline design across subjects including a reversal phase was used.

Baseline

Subjects recorded their measurements on the control chart during the assigned hour and left their completed charts at their workstations. Experimenters then collected the charts and completed checklists for each operator. Neither the operators nor the manager were informed of the operators' performance during this experimental phase.

Experimental Phase

The researcher provided the manager with index cards providing summaries of each operator's performance. The manager then presented the cards to the subjects approximately fifteen minutes after they had completed their measurements. Subject LB15 underwent two experimental phases separated by a reversal phase.

Reversal Phase

This phase was conducted with the same procedures as the baseline phase. Two of the three subjects (LB15 and LB12) were involved in the study long enough to perform under the reversal phase, while the third subject (Mill) did not perform under a reversal phase. The purpose of this phase was to further establish the relation between the intervention and subject performance. That is, it was expected that, after
performing at a higher-than-baseline level during the experimental phase, the subjects' performance would return to near baseline levels in the absence of the intervention.

**General**

The experimental conditions (phase) under which any one subject performed were determined by the stability of his own performance and that of the other operators under the current experimental conditions. In other words, a phase for one subject was changed only when his performance and that of the other subjects had stabilized in the phase under which they were currently performing. Stability was defined as performing at a level no more than plus or minus 20 percentage points from the average percentage points of the three immediately prior sessions. Due to his lack of performance stability, Mill worked under only two of the conditions in the study: baseline and an experimental phase. These stability criteria were established so that change in performance could not easily be accounted for by existing trends in performance or by factors such as experimental conditions being applied to other subjects.
CHAPTER III

RESULTS

Figure 3 shows the percentage of tasks completed correctly by the three machine operators who were subjects. Hereafter, the operator of the electronic lathe model number LB15 will be referred to as simply LB15, the operator of machine LB12 as LB12, and the operator of the electronic mill as Mill. The three graphs in the Figure are labeled accordingly.

During baseline LB15 correctly completed an average of 77.8 percent of the tasks assigned each day. This average was calculated by dividing the total number of tasks completed (35) during the nine days of data collection in this phase by the total number of tasks which could have been completed (45). The range of tasks completed was from 0 to 83.3 percent. During the feedback phase, this operator never performed below 100 percent of tasks completed correctly. This level of performance was maintained for 10 consecutive days. The reversal phase shows a change toward baseline performance on only one day. In the final feedback phase, again LB15 never performed at a level below 100 percent.

During the baseline phase, LB12 completed an average of 75.7 percent of the assigned tasks correctly. The range of tasks completed was from 0 to 100 percent. During the feedback phase, LB12, like LB15, never performed at a level below 100 percent over ten days of data collection. Unlike LB15, the reversal phase for LB12 showed a return to baseline levels; during the three days of this phase, LB12 averaged 66.7 percent tasks completed correctly. The range of performance was from 50 to 83.3 percent during the reversal, and the operator never performed at the 100 percent
Figure 3. Percentage of Checklist Tasks Completed Each Day for All Operators. Arrows indicate days when data were not available.
level in this phase. No additional feedback phase could be conducted with this operator due to his beginning work at another place of employment after day 35 of the study.

The mill operator performed at a higher level during the baseline phase than the other two operators. The average percent of tasks completed correctly was 84.2. This higher level of performance can be at least partially attributed to certain circumstances in the organization. For instance, this machine was typically assigned parts which did not require measurements as complex as those usually taken on parts manufactured by the LB15 and the LB12. Also, Mill tended to produce a larger number of parts per production run than the other two operators. This enabled him to become more accustomed to, and more accurate with, the measurements he performed on any one part. During the feedback phase, Mill performed at a 100 percent level with no variation.

The arrows on the graphs represent days when data were not available for a variety of reasons. In some cases, this situation was caused by an operator absence. At other times, a machine was being "set up" for a manufacturing run. Finally, no data could be collected on days when a machine was involved in manufacturing an experimental part. This was done periodically to determine if a machine was capable of manufacturing parts according to specifications prior to accepting an order from a potential customer.

In order to assess the reliability of observations by the primary observer, a second independent observer completed simultaneous observations for a portion of the sessions. Percent agreement between the two observers was calculated by comparing the items marked on the observers' checklists of operator tasks. An item was considered in agreement if both observers marked that item as being performed or not performed. Agreement was calculated on 8 of the 26 days for which data were
available for LB15 (30.7 percent of the measurement sessions). The average agreement for this measurement was 93.8 percent and ranged from 66.7 to 100 percent.

For LB12, agreement was calculated on 6 of the 25 days for which data were available (24 percent of the measurement sessions). The average agreement was 97.2 percent and ranged from 83.3 to 100 percent.

In Mill's case, agreement was assessed on 7 of the 29 days for which data were available (24.1 percent of the measurement sessions). Agreement averaged 95.3 percent and ranged from 66.7 to 100 percent.

Overall, interobserver agreement was assessed on 21 of the 80 occasions for which data were available. This is 26.3 percent of the measurement sessions. The overall agreement for all operators averaged 95.2 percent while ranging from 66.7 to 100 percent.

Table 1 presents the percent of measurements within control limits for the three machines across all parts made during the study. Although quality of work was not targeted by the study, this measure was tracked to ensure that emphasis on the checklist tasks did not lead to production problems.

The table shows data for two types of control charts. The first control limit was defined as 50% of the total tolerance allowed for the measurement on the blueprint of the part. The second, more stringent, limit was defined as 25% of the total tolerance allowed according to the blueprint.

Phase changes are denoted by broken lines in the body of the table. Corresponding indicators of the experimental phases are found in the far right column. "B" indicates the baseline phase, "F" indicates a phase in which feedback is present, and "R" indicates a reversal to baseline procedures. For the first subject, LB15, there is little change in measurements within 50% control limits across phases.
Table 1.
Percent Measurements Within 50% Control Limits and Within 25% Limits

<table>
<thead>
<tr>
<th>Sessions</th>
<th>LB15 50%</th>
<th>LB15 25%</th>
<th>LB12 50%</th>
<th>LB12 25%</th>
<th>MILL 50%</th>
<th>MILL 25%</th>
<th>PHASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>94</td>
<td>100</td>
<td>100</td>
<td>44</td>
<td>23</td>
<td>LB15-B</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>94</td>
<td>94</td>
<td>LB12-B</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>83</td>
<td>MILL-B</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>87</td>
<td>89</td>
<td>89</td>
<td>17</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>90</td>
<td>89</td>
<td>78</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>89</td>
<td>91</td>
<td>67</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>56</td>
<td>33</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>100</td>
<td>83</td>
<td>100</td>
<td>75</td>
<td>33</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>67</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>33</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>70</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>100</td>
<td>92</td>
<td>92</td>
<td>60</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>100</td>
<td>50</td>
<td>78</td>
<td>78</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>100</td>
<td>75</td>
<td>67</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>98</td>
<td>91</td>
<td>96</td>
<td>96</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>100</td>
<td>92</td>
<td>83</td>
<td>58</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>100</td>
<td>100</td>
<td>67</td>
<td>58</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>100</td>
<td>87</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

This operator's measurements of the critical dimensions fell between the control limits nearly 100 percent of the time through all four phases. The only deviation from this pattern occurred on the tenth day of the study when only 75 percent of the
measurements were within the control limits. The percentage of measurements within the 25% control limits taken by this subject was only slightly more variable. This measurement ranged from an average of 87 percent in the baseline phase to an average 96 percent in the final feedback phase.

This difference between baseline and feedback phases was more prominent for LB12. However, the trend was reversed. During the feedback phase, the measurements taken by this subject were within the 50% control limits only 78 percent of the time, while during the baseline and reversal phases, the measurements were within the limits 94 percent of the time. This difference also exists for measurements within the 25% control limits. During the feedback phase only 56 percent of the measurements met this criterion. During the nonfeedback phases the percentage was 75.7.

The results from Mill on this measurement resemble those of LB15. In the baseline phase, Mill averaged 72 percent of measurements taken within 50% control limits and 58 percent of the measurements within the 25% limits. During the feedback phase, the percentage of measurements taken by this subject which fell within the 50% limits averaged 100, and 88 percent of these measurements were within the 25% control limits. Just as for the measurement of tasks completed correctly, the results for Mill may be slightly misleading due to the short duration of the only feedback phase implemented.

In two cases, then (subjects LB15 and Mill), there is a small improvement in the proportion of measurements taken by the subjects which fall within control limits when feedback is present compared to when the feedback is not present. However, for the remaining subject, more measurements of critical dimensions were within set limits when formal feedback was absent than when it was present.
The results of the present study contain evidence that feedback procedures can be successfully used to increase the extent to which employees complete assigned work tasks. In the present case, overall success was shown even when the assigned tasks were fairly complex. Subject LB15 maintained a 100% rate of task completion during two separate feedback phases compared to lower rates of task completion during nonfeedback phases. Subject LB12 also maintained a perfect rate of task completion during feedback which represented a significant improvement over the rate during the nonfeedback phases. And, although Subject Mill often performed at a 100% task completion rate during baseline, performance never dropped below this level during feedback.

It is clear that these outcomes resulted from the feedback intervention and not extraneous environmental factors for several reasons. First, each operator exhibited abrupt increases in rate of task completion only upon, and immediately after, the introduction of the performance feedback. Second, the two subjects exposed to reversal phases showed slight decreases in their rate of task completion during this phase. The possible existence of carryover effects upon performance from the feedback phase merely adds to the evidence that changes in performance were due to the intervention rather than extraneous causes. Finally, an attempt was made to control for many other extraneous factors which may exist in the workplace. Care was taken to eliminate any suggestion that subjects should strive for a goal thus
minimizing the effects of goal-setting (Fellner & Sulzer-Azaroff, 1984) on the rate of task completion. Also, the researchers were not in direct contact with the subjects during the administration of feedback and so provided little if any social rewards to the operators contingent upon task completion. In addition, care was taken to document that the supervisor delivered feedback in a neutral manner so as to provide a minimum of social consequences contingent upon task completion. It is, however, possible that the operators shared their task completion rates with each other. This may have introduced some extraneous variables associated with competition or other social influences, but the feedback was delivered privately to each subject to minimize this occurrence as much as possible.

In addition to privacy, other dimensions of the feedback itself may have contributed to the success of the program. Duncan and Bruwelheide (1986) suggested that the source of the feedback, the mode of feedback transmission, and some aspects of the feedback message itself such as positiveness, accuracy, amount of information, timeliness, and specificity may influence the effectiveness of feedback. In the present study, the feedback was delivered by a supervisor whose recommendations to the company owner may influence such decisions as raises, shift assignments, and terminations. Also, the feedback was delivered in a form easily understood by the subjects (three short statements on an index card) by personal, rather than mechanical, means and contained 3 types of accurate, positive information. In addition, the delay between the completion of the tasks and the delivery of the feedback was minimal (about fifteen minutes). These feedback characteristics have also been suggested as ways of improving the behavior-controlling qualities of feedback by Balcazar, Hopkins, and Suarez (1986).
One characteristic of the feedback may have detracted from its effectiveness. The program employed summary data in the feedback rather than a description of the specific behaviors required of the subjects. This lack of specificity has been cited as a characteristic of feedback with less than optimal effectiveness (Balcazar, Hopkins, & Suarez, 1986; Duncan & Bruwelheide, 1986). However, no attempt was made in the present study to empirically compare the effects of various feedback characteristics, and this endeavor will be left to future research.

The present intervention can also be considered a success in reference to factors other than the increase of task completion. One of these factors is the time and training demands placed upon the organization. Each operator spent approximately one hour learning how to plot data points upon the graphs. The supervisor was trained to deliver feedback appropriately in one-half hour. Also, the sampling procedures were streamlined so that the operators were required to spend a minimum of time engaged in this activity. The supervisor's time contributions were similarly minimal in that he spent approximately five minutes per day actually delivering feedback. It is recognized that, in the absence of researchers, an organization would be required to develop its own control charts, task checklists, and mode of feedback. However, the present program serves as an example of a quality program which yields positive results concerning worker performance while placing minimal demands upon training and other time-related demands of an organization.

Unfortunately, the results concerning the overall quality of outputs (i.e. the percent of parts within control limits) were not as positive. The quality of parts produced did not seem to differ significantly dependent upon the presence or absence of performance feedback. That is, those parts which were high in quality during baseline remained so during feedback, those low in quality remained low during
feedback, and those parts produced which were highly variable with respect to quality during baseline, retained this variability during feedback.

For two of the subjects, the machine used may be a contributing factor to these results. LB15 used a new, highly advanced, "top-of-the-line", electronic lathe. Therefore, the parts produced by this machine may have been of high quality regardless of the task completion rate of its operator. On the other end of the scale, the drill press used by Mill was an older machine which the supervisor acknowledged produced parts with a high degree of variability in quality. This too occurred without regard to the rate of task completion of the operator.

The lack of improvement in the quality of products produced by LB12 when higher percentages of assigned tasks were completed is more difficult to explain. It may be that, during the feedback phase, this subject was producing a part with characteristics, such as hardness of the metal or relatively tight control limits, which led to a decrease in quality measures. Although no data regarding such characteristics were collected, their existence has been noted in other quality control programs. Deming (1975) referred to such unidentifiable and, hence, uncontrollable causes of quality variation as random causes. Those causes of quality variation which could be identified and positively affected, such as task-relevant behaviors, were referred to as assignable causes. A high degree of consistency in both types of machines used and types of parts produced may eliminate some of the random causes of quality variation, and so may enable future researchers to accurately identify and influence assignable causes of variation.

The trend shown in these results indicating that contingent feedback can be used to maintain fairly complex employee behaviors such as those required in the implementation of a statistical quality control program should be welcome news to
behavior analysts. One implication is that feedback procedures can be used to combat the possible employee resistance documented by Levi and Mainstone (1987) to completing tasks such as plotting data points and analyzing summary data.

Behavior analysts should also be encouraged by these results to administer feedback programs which are designed to address a wide range of other complex behaviors similar to those performed by the machine operators. Classes of behavior which may be positively affected by such feedback programs include the work performed by accountants, engineers, economists, draftsmen, and students. All of these vocations require the accurate use of complex math skills, the strict adherence to program or manual guidelines, and, quite commonly, the employment of small handheld tools. In addition, any profession in a service industry which requires the sampling of products or other outcomes through the use of physical measurements or even some type of opinion poll may benefit from the application of feedback procedures similar to those used in the present study. Such professions may include real estate and stock broking, poll taking, advertising, and even teaching.
APPENDICES
Appendix A

Informed Consent Form
INFORMED CONSENT

THE EFFECTS OF PERFORMANCE FEEDBACK ON THE IMPLEMENTATION OF A STATISTICALLY-BASED QUALITY CONTROL PROGRAM

I am willing to participate in the research project being conducted by Gordon Henry and Dr. William Redmon. I understand that this study will examine the effects of supervisory feedback on the completion of implementation tasks associated with establishing a quality control program. I also understand that: my participation is completely voluntary, my performance data will remain confidential, and that neither my decision to participate or not to participate nor information gained during this study will affect my compensation or my current employment status.

Date: ____________________

Name (please print your complete name): _____________________________

Signature: ___________________________________
Appendix B

Operator Sampling Plan
Appendix C

Human Subjects Institutional Review Board Approval
TO: Gordon O. Henry
FROM: Ellen Page-Robin, Chair
RE: Research Protocol
DATE: March 13, 1988

This letter will serve as confirmation that your research protocol "Establishment of Implementation Behaviors in a Quality Control Program" is now complete and has been signed off by the HSRB.

If you have any questions, please contact me at 387-2547.
BIBLIOGRAPHY


