

Individual and group productivity enhancement in a service department

Citation for published version (APA):

Kleingeld, P. A. M., & Tuijl, van, H. F. J. M. (1992). *Individual and group productivity enhancement in a service department*. Eindhoven University of Technology.

Document status and date:

Published: 01/01/1992

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Individual and group productivity enhancement in a service department

P.A.M. Kleingeld
H.F.J.M. van Tuijl

ARW 02 TUE

Department of Industrial Engineering and Management Science
Eindhoven University of Technology

June, 1992

Individual and group productivity enhancement in a service department

INTRODUCTION

This paper focuses on the development and implementation of ProMES for about 300 technicians servicing photo copiers at client offices throughout the Netherlands. The case illustrates some of the complexities of real life work situations ProMES can handle. Three points will be dealt with in particular. First, because service technicians in this case were in some respects dependent on their colleagues, but in other respects not, the system developed was partly based on individual performance data and partly on data gathered at the group level. Secondly, the differences between the photo copiers serviced by the department could only be dealt with adequately by a rather complex ProMES system: separate sets of contingencies had to be developed for each type of photo copiers. Thirdly, the relationship between ProMES and performance appraisal and reward issues had to be dealt with. The main reason was that in this case ProMES feedback reports contained precise information on the performance of individuals. Supervisors got copies of the feedback reports, i.e. they knew exactly how well each technician was doing his/her job. This knowledge could not be ignored in the annual performance appraisal assessments. In addition, performance appraisal scores influenced the amount of bonus received by technicians at the end of the year. So, the issue of how to deal with ProMES scores in the context of performance appraisal and rewards had to be settled. At several places, we will refer to these three points.

We will successively describe some *context information*: the company, the immediate motivation for installing ProMES; *the procedure* followed in developing ProMES, *the system* and its implementation, the *preliminary results*, and finally, *some adaptations* of the system instigated by those preliminary results.

SOME CONTEXT

ProMES was developed at the field service department of a company selling and servicing office equipment (faxes, laser printers and photo copiers) in the Netherlands. The service department was chosen because of the management's intention to make service the company's competitive edge. Probably for the same reason, the management wanted to install a payment by results system in the service department, as they already had done in the company's sales department. ProMES could provide the performance data necessary for such a system. A final consideration was the technicians' recurring request for feedback. A large amount of performance data were gathered already by the company, but

very few of these data were fed back to the technicians. Because the vast majority of service resources were spent on repairing and maintaining photocopiers, it was decided to leave faxes and laser printers out of the ProMES to be developed.

The service department was organized geographically. There were eleven geographical regions at the start of the project and there are fourteen now. Each was serviced by a group of some twenty technicians headed by a supervisor. The technician's job consisted of repairing machine malfunctions and doing preventive maintenance. The maintenance was done at the same time as the repairs, in that there was no prearranged schedule for preventive maintenance. The regular course of action ran more or less as follows: the client makes a phone call to the company, indicating that his photo copier is out of order. A first attempt is made to solve the problem by giving the client some instructions on the phone. If this does not succeed, the planning department instructs a technician to go to the client, repair the machine and do the preventive maintenance suitable at that time, given the maintenance history of the machine. The technician drives to the client's location, repairs the machine, follows the prescriptions with regard to preventive maintenance, fills in a card containing information on the machine's repair history, and leaves for the next client. Before he does so, he makes a phone call to the home office, informing his dispatcher about: the time he took to travel to the customer, the time needed to repair and maintain the photo copier, the spare parts he used, and the number of copies on the counter. He also orders additional spare parts as he runs out of stock. (The supplies asked for are delivered at the technician's home address by a special service.) The dispatcher tells him his next job, and off he goes. It is important to know that his next job can be at any client in his region: the planning department allocates technicians to jobs in such a way that agreements with clients regarding service within a certain time limit are fulfilled. For that reason, there are no fixed connections between technicians and clients: any technician can be allotted a particular repair, provided that he has the requisite knowledge. It is the planning department's responsibility to use the available technicians efficiently and at the same time to fulfil the company's obligations to the client.

To list the most salient aspects of the technician's job: The job is a rather individual one. The technician works on his own, meets his colleagues only at monthly meetings, and his supervisor at most one or two times per month during a repair visit for a client. He is primarily dependent on his own knowledge and skill, but he can ask for help from a product support department, when unable to solve a technical repair problem on his own. To a certain degree he is dependent on others too: he has a hard time repairing and maintaining a machine serviced at a prior occasion by a colleague who did a bad job: in such a case he has to invest a lot of time and spare parts to make good the damage caused by his colleague. The way a photo copier is handled by a client is important too, because it influences the occurrence of malfunctions. The timeliness with which the warehouse fulfills his requests for spare parts influences the technician's stock of parts he

carries in his car, and in turn, the percentage of repairs that cannot be finished adequately because of lack of spare parts. In such cases a return visit is necessary.

Some of the company's characteristics are quite relevant for understanding what happened during the process of development and implementation. First of all, there was a lack of communication both horizontally and vertically, caused by the nature of the service work to be done. Technicians had very few opportunities to communicate with one another and with their management: their supervisor and the levels above the supervisor. Because of this lack of communication, all kinds of prejudices arose and continued to exist. Secondly, the company's culture could be characterized as top down: decisions were made at headquarters; technicians should do as they were told. In addition, there were large differences in culture between the sales and the service department, and between headquarters and the field: sales and headquarters could be characterized as a white collar culture, field service as a blue collar culture. In addition to that, feelings of inequity existed in the field service regions because sales reps and managers could earn significant incentives, whereas service technicians could not.

PROCEDURE

The company had decided to develop ProMES for the service technicians of the field service department for the above mentioned reasons. In addition, it was decided to first develop and implement the system in two regions, and to install the system developed by the two trial regions later in the remaining regions, provided that the implementation in the two trial regions had been successful. The two regions were selected on several grounds. One was an urban region: Utrecht, the other a rural one: East. One was cooperative, the other one critical with regard to a system like ProMES. Next it was decided to develop ProMES in each of the two regions separately, in monthly meetings during paid overtime hours, with design teams consisting of all the technicians of the region (about 20 people, all male), the region's supervisor and two facilitators (this paper's authors). This choice necessitated several inter-region meetings in which consensus between the two regions had to be reached before elements of the system could be discussed with the management. The decision to work with all the technicians in the design team was made because of the highly isolated nature of the technician's job. Communication between job incumbents during working time simply was impossible. This decision resulted in extremely large design teams. To structure the discussions in these teams, several methods (to be discussed in more detail later) were used in addition to the normal interactive group discussion, such as nominal group technique and discussion in subgroups (at monthly meetings) and delphi method (to stimulate between meeting information exchange). ProMES was first introduced to the supervisors of all regions and their consent was obtained. Next, the two trial regions were informed. The

reactions of the critical region were such that the management was urged to give a written guarantee that the system to be developed would not be used for performance appraisal or reward purposes until all parties concerned (i.e. both management and technicians) agreed that it would be a good thing to do so. After this guarantee was given, the development process was executed according to plan.

SYSTEM DEVELOPMENT

In each region, it took 13 meetings of about 2.5 to 3 hours each to develop the system. Because there were no meetings during holiday periods, the whole development process took one year and a half, from June 1989 till September 1990. During this period, two students spent each, one after the other, eight months full time on the project to fulfil their master thesis' obligations. Their job was to prepare the meetings, to act as co-facilitators, to adapt the company's computerized information system to produce ProMES reports, to do all the paper work involved (reports, the delphi material, and so on).

The following products and indicators were developed.

Product 1. Quality: to repair and maintain photo copiers as effectively as possible.

Indicator 1.1 Mean copies between calls: the average number of copies made between a technician's repair visits and the first malfunctions after those visits.

Indicator 1.2 Percentage repeat calls: the percentage of repair visits made by a technician within 5 working days of the original repair (because of inadequate repair).

Indicator 1.3 Compliance with preventive maintenance prescriptions: percentage of preventive maintenance procedures correctly followed.

Product 2. Cost: to repair and maintain photo copiers as efficiently as possible.

Indicator 2.1 Parts cost per call: the average amount of money spent on replacement parts.

Indicator 2.2 Labor time per call: the average amount of time used for repairing and maintaining photo copiers.

Indicator 2.3 Percentage return calls due to car stock: the percentage of visits caused by a lack of spare parts in the technician's car stock.

Product 3. Administration: to keep records of repair and maintenance activities as accurately as possible.

Indicator 3.1 Accuracy of history card: percentage of required information filled in correctly.

Indicator 3.2 Completeness of claims: percentage of potential claims (parts replaced within their warranty period) submitted.

Product 4. Attendance: to spend the available-time on work related activities.

Indicator 4.1 Percentage of capacity used: the percentage of labor contract hours actually spent on the job.

Product 5. Ambassadorship: to behave as correctly as possible on the job.

Indicator 5.1 Correctness of behavior: the percentage of important social behaviors demonstrated on the job.

Products and indicators were generated by the design teams according to the procedures described above. During the discussions several products and indicators were rejected. To give an example, the proposed product 'paper sales' (technicians were supposed to turn the client's attention to the fact that the company was selling paper too) was rejected, because technicians considered this an activity which had nothing to do with their core responsibilities (repairing and maintaining photo copiers) on which they should spend their time and effort. When this issue was discussed with the management, the decision to exclude that product from the system was confirmed. The discussions with the management of the field service department were intense and valuable. This was partially because for many technicians, this was the first real contact with the managers. Management expressed their appreciation for the work done by the design teams. Several important issues were settled in direct interaction. For example, management insisted on technicians wearing official company work clothes, but the work clothes were not delivered on time, so that technicians often had to do their job dressed unofficially. A similar problem had to do with the delivery of spare parts. Management asked for low percentages 'return calls for car stock', but the warehouse made many mistakes in supplying the technicians. It was not possible to resolve all these issues, but they were brought to the attention of management and taken seriously by them. The controllability criterion which products and indicators should fulfill was discussed here. Responsibility was only accepted by the technicians in cases where solutions could be given or promised for the near future. One general solution offered was to use running averages over larger periods as measures in addition to scores per month. By taking averages, the influence of chance factors resulting in accidentally high or low scores is minimized, thereby giving a more accurate picture of a stable 'true' score.

It should be mentioned that the discussion with the management also resulted in the addition of an indicator. Indicator 1.3, 'compliance with preventive maintenance procedures', was not in the list generated by the design teams. In discussing the products and indicators proposed to management. Management felt that improving performance on the proposed indicators could be beneficial to the company on the short run, but disadvantageous in the long term. The reason was that preventive maintenance could easily be neglected in favor of labor time and parts' cost per call. The net short term effect would be high scores on these indicators and acceptable scores on 'mean copies between calls' and on 'percentage repeat calls', both short term quality indicators. A probable long

term effect would be a decrease in the general condition of the photo copiers, eventually resulting in high machine malfunctions. To prevent this, a new indicator was developed which measured preventive maintenance quality. The indicator consisted of a large checklist covering the main elements of the preventive maintenance procedures for each type of photo copier at several points in the machine's history. It was agreed that supervisors would check the preventive maintenance done by technicians by sampling. Because of the impracticality of taking large samples, it was decided that this indicator would result in a score for the region, not for individual technicians. The same decision was made with regard to 'accuracy of history card' and 'correctness of behavior', both indicators measured by means of a checklist filled in by the supervisor on a sample of technicians' visits.

Apart from the impracticality of gathering individual performance data on the checklist indicators continuously, or by taking large samples, there was another reason to deal with these indicators at region level. As mentioned above, technicians were to a large degree dependent on the quality of repair and maintenance done by their colleagues. If every technician does proper preventive maintenance and correctly records repairs done, parts replaced, etc., colleagues who have to do the next call are in a comparable starting position. That is to say, individual score's on indicators such as 'labor time', 'parts cost', 'mean copies between calls' and 'percentage repeat calls' are only valid indicators for individual performance if all technicians correctly execute the preventive maintenance and administration part of their job. So, group commitment with regard to decisions on these indicators is needed.

The four indicators 'compliance with preventive maintenance procedures', 'accuracy of history card', 'correctness of behavior', and 'percentage of capacity used' were newly developed indicators by the design team. The company already gathered information on the other indicators, but as we shall see below, there were doubts on the validity and meaningfulness of that information. The main reason for the latter point was that data concerning different types of photo copiers, i.e. high volume, middle volume and low volume machines, were not comparable. Even though this was known, the company simply added the scores across these types of copiers, resulting in feedback that was not meaningful. As will be shown below, the contingency technique offered by ProMES can solve this problem: seemingly incomparable scores on identical indicators can be added in the same way as score's on different indicators.

OPERATIONALIZING THE INDICATORS

The operationalization of the indicators turned out to be a very time consuming and complex activity. In between development meetings, the facilitators worked on the indicators with the help of product support specialists, supervisors, and a programmer. The results were reviewed by the design team during development meetings.

LEVEL OF MEASUREMENT

An objective of this ProMES system was to measure the productivity of service technicians both as individuals and as a group. The indicators under the control of individual technicians should be measured individually. For three indicators however, it turned out to be too costly and time-consuming to do so. For that reason, it was decided to measure individual performance by sampling and to use the sampled scores as scores for the whole group.

TYPE DEPENDENCY

It was found that 5 indicators were 'type-dependent'. That is, technical characteristics of each type of copier determined to a large extent the possible performance levels on these indicators. For example, historical data showed that on model X a Mean Copies Between Calls of 15,000 is an excellent result, whereas on model Y it is just average, and on model Z it is very bad. This has far-reaching consequences for the development of contingencies, because, for these 5 indicators, a separate set of contingencies had to be developed for each of the 26 types of copiers included in the system.

Some indicators turned out to be similar to measures already in use in the service department. However, most existing measures were not suited for measuring productivity on an individual level, did not take into account differences between types of copiers, or included elements the technicians could not control. The existence of these measures meant that much of the data needed for the ProMES indicators was already being reported by the service technicians on a daily basis (e.g. how long did the visit take, which spare parts were used). Hence, the operationalization focused on defining the indicators and combining the data. All definitions refer to a measurement period of one month, which was customary in the service department.

There were ten indicators in all, three of which are measured type-independently at group level (Compliance with maintenance procedure, Accuracy of History Card, and Correct Behavior). The other seven indicators were measured individually. Five of these were type-dependent (MCBC, percent repeat, labor time per call, parts cost per call, and percentage return calls due to car stock). The remaining two indicators, completeness of claims and percentage of capacity used, were type-independent.

DIFFICULTIES IN OPERATIONALIZING THE INDICATORS

For a few indicators, the operationalization process was rather straight-forward. Most indicators, however, posed serious difficulties, three of which will be discussed.

The first problem is related to the existence of *return calls*, visit(s) following the initial visit that are necessary to complete repairs and maintenance. These occur because

the technician lacks spare parts, or the time or knowledge to complete the initial call. In about 10 percent of the cases, the return call is handled by another technician, which complicates the allocation of labor time and spare parts usage to the individual technicians involved. Keeping in mind the objective of measuring individual productivity, it was decided to exclude calls involving more than one technician from the set of calls that is used to calculate parts cost per call and labor time per call. In that way the results of individual technicians are not influenced by their colleagues.

The second complicating factor is a time-lag that sometimes exists before indicator values can be calculated. An example is *Mean Copies Between Calls (MCBC)*. In order to get this measure, the counters are compared for two consecutive calls on a machine, the difference being the number of copies made between the two calls. This is attributed to the technician involved in the first of the two calls. Averaged across calls, the MCBC is calculated per type of copier. So, it is not possible to calculate the number of copies between calls until the next call has occurred. This makes it difficult to calculate the MCBC for a technician at the end of a month, because it is very likely that not all of his repairs in that month have resulted in a new call. In fact, before the end of the month, only the less successful repairs will have resulted in new calls, the MCBC of which would not be an accurate measure of the technician's performance. For this reason, the design team decided to attribute MCBC results to the month in which they can be calculated. This is the only way to arrive at complete information, although part of the MCBC results originated from months prior to the feedback period.

Thirdly, it was found that there were no existing measures available for the indicators "Compliance with maintenance", "Accuracy of History Card" and "Correct Behavior". Although some technicians were skeptical about introducing subjectivity into the system, the consensus eventually was that the region supervisor would be the person most capable of measuring to what extent the technicians met the demands posed by these indicators. He could do that on two occasions: during a visit, by watching the technician at work, or after a visit by looking at the results of the work done by a technician the previous day. Yet, on practical grounds it would be impossible to get an accurate monthly measure of all 20 technicians' individual performance. This is unfortunate because two of these indicators reflect the interdependence between the technicians and could be used to insure the validity of performance data on the individual indicators. For example, a technician could obtain a high effectiveness score on labor time per call by carrying out only part of the maintenance procedure. This must be avoided, because it negatively influences the long-term quality of the machine. It was decided to measure the above indicators by sampling, resulting in group level indicator values that are allocated to the region as a whole and to all individual technicians. This again emphasizes the fact that the interdependencies between technicians are mainly reflected in the indicators "Compliance with maintenance procedure" and "Accuracy of History Card" and that there is a joint responsibility to perform well on these indicators.

Most of the work on these checklists was done by the facilitators. Critical incident interviews were conducted with supervisors, field service management and product support specialists. These resulted in two checklists per indicator to be used during a visit and after a visit respectively, and a procedure for applying these checklists.

EVALUATION OF THIS PHASE

The operationalization process was characterized by a general lack of involvement of the design team. Most major decisions were made during the development meetings, but the facilitators rather than the technicians did most of the thinking and deciding. Although this was the only procedure possible, there was a risk of causing limited insight into the system, low perceived validity and, as a consequence, low commitment on the part of technicians and supervisors.

There indeed was a temporary lack of insight into the system: during the last development meeting and the first feedback meetings, the facilitators were called upon to explain how certain indicators were operationalized. Most doubts about the validity of the indicators were taken away when the facilitators explained why certain choices were made. An important limitation in this respect is the fact that most indicators are not 100 percent controllable by individual technicians. The indicators were operationalized in such a way as to exclude as many as possible uncontrollable factors from the indicators, without diminishing the relevance of the indicator. The technicians agreed with most choices that were made. Some minor changes were suggested and incorporated into the system.

All in all, the indicators were operationalized satisfactory. However, problems did occur in gathering data in accordance with these operationalizations. These are discussed together with the feedback meetings.

CONTINGENCIES

The design of this stage differed in some respects from the usual approach (Pritchard, 1990). First, the first two stages in the development of the contingencies ('identifying maximums, minimums, and zero points' and 'establishing maximum and minimum effectiveness scores') were switched. Secondly, scaling techniques were used extensively in the initial stages of establishing maximum and minimum effectiveness scores. Thirdly, historical data were used to establish minima, maxima, and zero points for the type dependent indicators. These differences will be discussed.

ESTABLISHING MAXIMUM AND MINIMUM EFFECTIVENESS SCORES

Paired Comparisons. To get a first impression of the relative importance of the indicators the technique of paired comparisons (Edwards, 1957) was used. In applying this technique, the individual technicians were asked the following question for each pair of indicators: "Suppose you perform at the expected level (not good/not bad) on all indicators; on which of these two indicators would you want to perform maximally to maximize the value of your performance for the organization?". The results of these paired comparisons were converted into a ratio scale for maximum effectiveness scores by giving the maximum with the highest importance an effectiveness score of +100 and deciding on the effectiveness score of the least important maximum (close to but above zero). The latter effectiveness score was determined by asking how effective a maximum value on this indicator would be, compared to the effectiveness of +100 for the most important maximum. The effectiveness scores in between are determined by the results of the paired comparisons. The minimum effectiveness scores were determined in an analogous way, the effectiveness score of the minimum with the highest importance being determined by the design team.

Ranking. A ranking technique was used in both regions, in addition to the paired comparisons. Each individual technician was asked the following question: "Suppose you perform at the expected level on all indicators. What is the first indicator on which you would want to perform excellently to maximize the value of your performance for the organization? What is the second one, third one etc.?" An analogous question was asked for the minimum effectiveness scores. By averaging the rankings across technicians a ranking of maximum and minimum effectiveness scores was calculated.

Discussion. The initial effectiveness scores determined by paired comparisons and ranking were the starting point for group discussion on the exact positioning of minimum and maximum effectiveness scores. Two additional criteria were used in this discussion:

1: The relative importance of the products should be correctly reflected by the indicators.

This criterion was used because not all indicators are completely independent. For example, there is a dependence between Mean Copies Between Calls and percent repeat calls: if a technician succeeds in lowering his percent repeat calls then his MCBC will probably be higher. Consequently, the product being measured (in this case quality) might be overweighed in the measurement system.

The importance of each indicator was determined by taking the range from minimum to maximum effectiveness score. This is a better way of establishing the relative importance of indicators than comparing the maximum scores, provided the minimum effectiveness scores too are determined with relative certainty. In addition, problems with determining the importance of indicators having a maximum effectiveness score of zero or with indicators having asymmetrical effectiveness scores (e.g. -65, +40) are avoided using this approach.

2: *Maximum effectiveness on quality should compensate for minimum effectiveness on cost and vice versa.*

The design team felt that there was a trade-off between the two products quality and cost. This trade-off should be reflected in the maximum and minimum effectiveness scores of the indicators belonging to those products. So a technician with maximum (minimum) effectiveness on quality and minimum (maximum) effectiveness on cost should have a total effectiveness score of about zero, indicating his performance is about average.

IDENTIFYING MAXIMUMS, MINIMUMS, AND ZERO POINTS

During the development of the indicators it became clear that the design team would not be able to identify maximums, minimums, and zero points for most of the type-dependent indicators. Hence, it was decided to generate historical data to help identify these points. The following decisions were made in generating the data:

- * The data should be collected per type of copier.
- * A minimum monthly number of 5 calls per technician per type of copier was used to avoid atypical results. For example, a technician performing a single call on type X might have either 0 percent or 100 percent repeat calls. Neither indicator value would be very informative.
- * The data should cover a one year period to make sure any seasonal trends are incorporated into the contingencies.

For all type of copiers a frequency distribution of indicator values was generated for all five type-dependent indicators.

In a discussion between management and the design teams, management stated that, in their opinion, the service technicians had performed according to expectations during the period concerned. Therefore, they suggested that the average national indicator values on all types of copiers be considered the expected level of performance (not good/not bad). The design teams agreed with this point of view.

The facilitators suggested that the range of results in the frequency distribution might be used to determine minimum and maximum indicator values. Management and the design team agreed, indicating that this would be the only practical way to determine these values.

It has to be noted that it would probably be incorrect to base contingencies on historical data without having management and design team ask themselves whether the resulting maximums, zero points and minimums, which are of a descriptive nature, can be used normatively. In the worst case this might result in contingencies formalizing ineffective policies. In this organization, the service department used procedures and working methods that were quite sophisticated when compared to many companies in the industry, so there was no risk of reinforcing ineffective policy.

For the 26 types of copiers, minimum, maximum and zero points were established for the five type-dependent indicators, using historical data. For the remaining 50 types (constituting less than 10 percent of all calls), no accurate values could be established due to the very small amount of historical data. These types were excluded from the system.

The maximum, minimum, and zero points of the five type-independent indicators were determined through discussion, as is usually done in a PromES development.

POSSIBLE CONSEQUENCES OF SWITCHING STAGES

The first two stages in the development of the contingencies (identifying maximums, minimums, and zero points and establishing maximum and minimum effectiveness values) were switched to bridge the time needed to collect the historical data to be used to establish minima, zero points and maxima for the type-dependent indicators. Two possible pitfalls were identified.

First, this sequence does not take into account indicators which have a maximum effectiveness score of zero (i.e., equal to the expected level). In this case, a few indicators might have had such contingencies, e.g. Compliance with maintenance procedure. Yet, when the technicians attention was called to this, they held the view that they should have the opportunity to 'score some points' on each indicator, i.e. that every maximum should have a positive effectiveness score.

Secondly, this procedure also neglects possible differences between types of copiers concerning the relative importance of indicators, requiring different strategies. For example, type X may be a relatively labor-intensive type, where labor time is a very important cost indicator, whereas type Y may be a relatively parts-intensive type where parts cost per call is a very important cost-indicator. These differences would not be reflected in the system, because the maximum and minimum effectiveness scores of the indicators are the same for all types of copiers. Although these differences between types exist, it was decided by facilitators and management not to take them into account, because it was felt that the resulting complexities could not be handled.

DRAWING THE COMPLETE CONTINGENCIES

The objective of this stage is completing the contingencies by filling in the remainder of the points. The total number of contingencies was 135, consisting of 130 type-dependent contingencies (5 indicators for 26 types of copier) and 5 type independent contingencies.

In determining the precise form of the contingencies for the type-dependent indicators, a well-known type was used as an example. Using transparencies, the facilitators started by simply drawing straight lines between minimum and zero point and between zero point and maximum. This caused an inflection point at the zero point for all indicators, which according to the design team would not correctly reflect the change in effectiveness

going from an indicator value just below to one just above the zero point. This change in effectiveness should, in their opinion, be about linear. Keeping in mind that most technicians generally have indicator values around the zero point, it was decided to draw contingencies which were relatively steep around the zero point and relatively flat near the minimum and maximum. The exact location of the resulting inflection points was determined using the frequency distribution of historical indicator values. The design team decided that 10 percent of the observations should be above the upper inflection point and that 10 percent of the observations should be below the lower inflection point.

After the design team had developed the completed contingencies for this type of copier, the facilitators were authorized to develop the contingencies for the remaining 25 types of copier, applying the rules mentioned above. Figure 1 (page 14) shows two type-dependent contingencies (Mean Copies Between Calls and percent repeat calls) for two types of copiers (type 4100 and type 7150). Note that the shape of the contingencies is roughly the same for the two types. However, the type of copier determines to a large extent which indicator values correspond to minimum, zero and maximum effectiveness.

The type independent contingencies were determined through discussion.

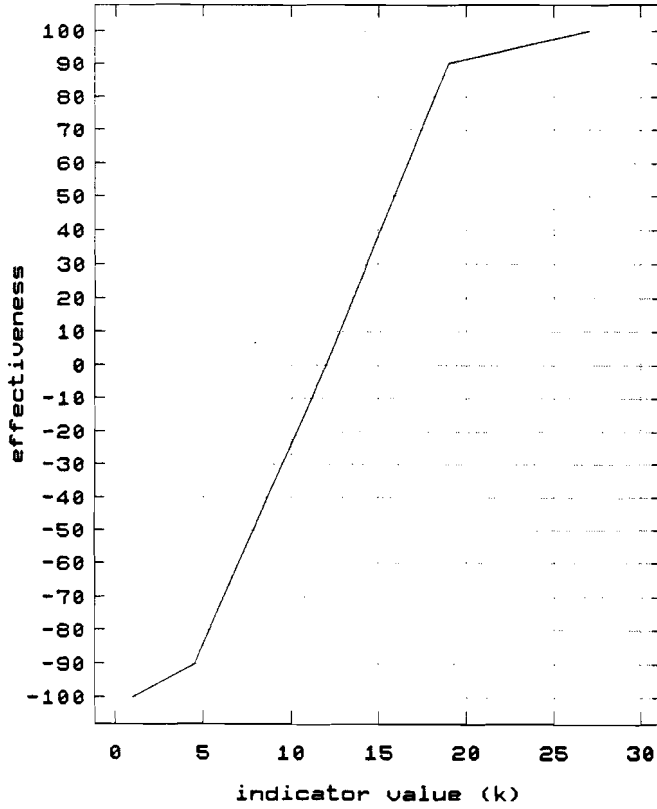
FEEDBACK REPORTS

BACKGROUND

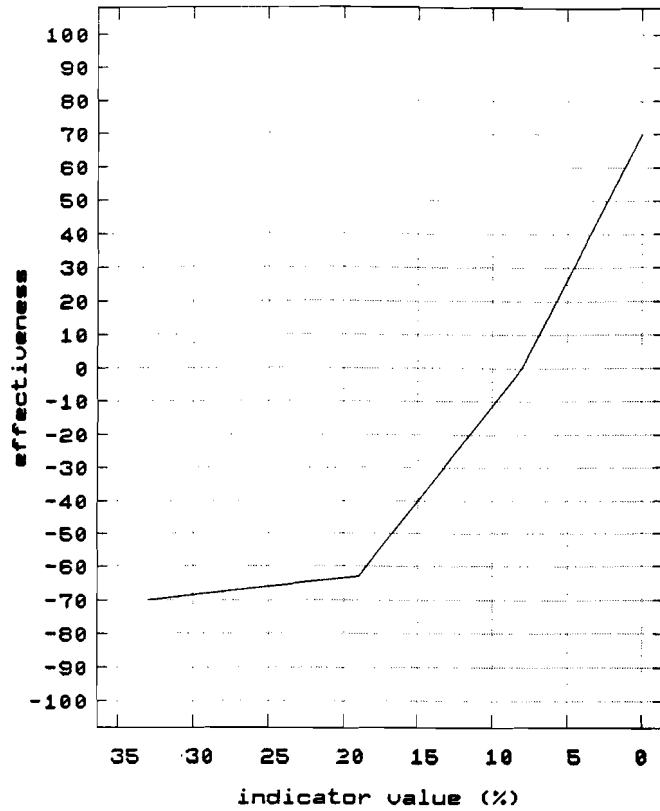
Generally, in designing feedback systems, interdependency between group members has to be taken into account. The ProMES system for these technicians is aimed at employees who work individually, but are dependent on their colleagues. According to Matsui et al. (1987), in a case of high interdependence between group members, effectiveness of task feedback is maximized when it provides information on individual and group performance. Just providing group feedback will not cause those individuals who are below target to improve their performance, as long as the group is on target. Also, possibilities for social loafing are created. Just providing individual feedback will not cause individuals who are on target to improve their performance when group performance is below target. Mitchell and Silver (1990) argue that setting individual goals in an interdependent task results in lower performance than setting group goals, group goals plus individual goals, or no goals at all because of dysfunctional effects of competitive feelings, strategies and behavior resulting from individual goals.

Keeping in mind the interdependence between technicians, it is clear that there should not only be an individual feedback report, but also a group feedback report. From a practical point of view, a group feedback report is indispensable as a common frame of reference and a means to stimulate mutual help to improve the effectiveness of individual technicians and thereby the effectiveness of a group as a whole.

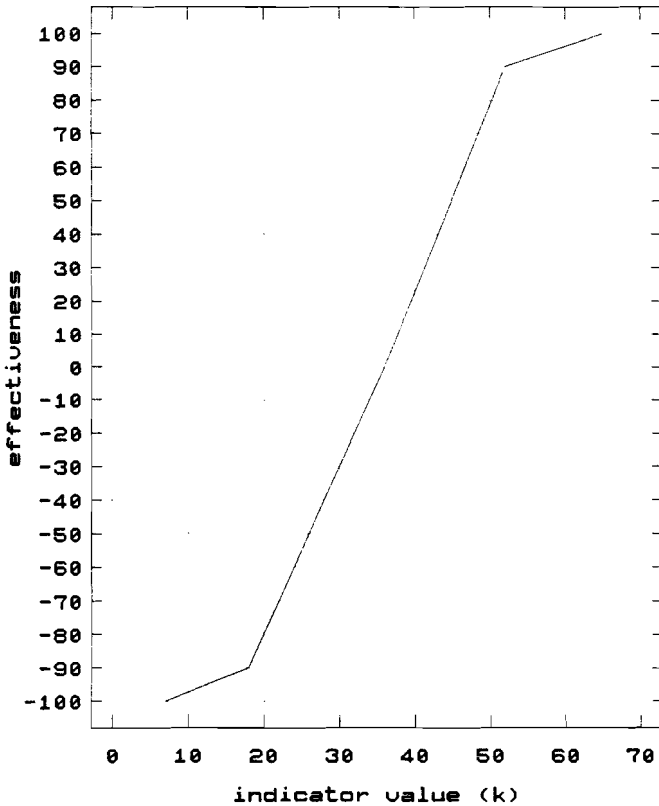
Type 4100 Quality
MEAN COPIES BETWEEN CALLS



Type 4100 Quality
PERCENT REPEATCALLS



Type 7150 Quality
MEAN COPIES BETWEEN CALLS



Type 7150 Quality
PERCENT REPEAT CALLS

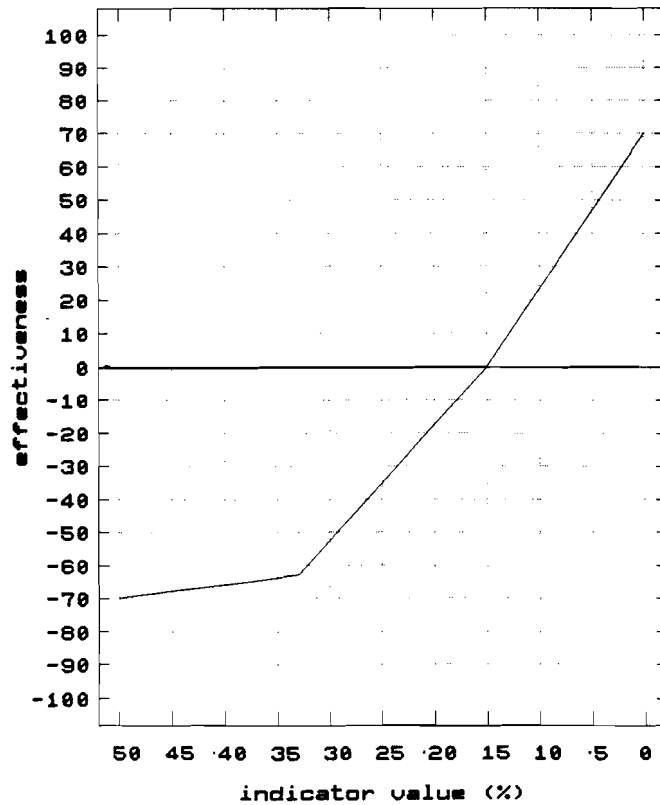


Figure 1: 2 contingencies type 4100 and 2 contingencies type 7150

PROCESS

As a basis for discussion in the design team, the facilitators prepared a first draft of an individual and a group feedback report, based on their ideas on what would be useful feedback and how it might be presented. Three questions were posed by the facilitators, dealing with the feedback report, feedback meetings, and the extent to which the reports should be private or public.

1. *"What information on your own performance and on the performance of others would you like to have ?"*

After some discussion, there was agreement that every technician should have detailed information on his own performance. In addition, every technician should have information on group performance.

2. *"How would you like to use the reports during the feedback meetings ?"*

Initially, there was difference of opinion in the design team on how to use the feedback reports. Some technicians argued that every technician should judge for him/herself how to use the feedback reports. Only when the individual effectiveness score compares unfavorably to the region effectiveness score, should a technician try to improve. Other technicians argued that communication and exchange of information between technicians would be essential to improve the effectiveness of individual technicians and thereby the region effectiveness, especially for the type-dependent indicators. The knowledge and experience of technicians who perform well on these indicators should be shared with technicians who experience problems on these indicators. After ample discussion the former technicians were convinced by the latter. In this discussion the facilitators stressed that a feedback meeting should be a factfinding exercise, not a search for excuses or a method for attaching blame. This might have put some technicians at ease who still appeared to have some fears concerning objectives and use of the feedback meetings. Finally, the design team reached agreement on the following general procedure during the feedback meetings:

- a. Type independent indicators: discussion in full session, conducted by the supervisor. The supervisor should, among other things, present his findings pertaining to the region indicator checklists.
 - b. Type-dependent indicators: discussion in small "type" groups (about 5 technicians). The groups should be discussing one or two types of copier on which all members receive feedback. Information per type of copier, especially information on individual differences, would be necessary in order to have fruitful discussions on causes of performance differences between technicians.
 - c. Plenary announcement of the decisions and strategies arrived at in the "type" groups and written confirmation of decisions and points for attention in the next month(s).
3. *"Who should get (parts of) the feedback reports ?"*

There was agreement in the design team that the region supervisor should have access to all individual information. The general opinion was that the supervisor is responsible for the functioning of the technicians in his region. He should consequently be able to approach technicians individually in all matters regarding the functioning of individual technicians. The service management should receive region feedback reports, because they are responsible for the effectiveness of the regions as a whole.

A revised draft of the feedback reports, approved by the design team, was presented to management. Management agreed with most of the design team's suggestions on how to use the feedback reports. However, management objected to including detailed information per type of copier in the reports, because this would take too much time to program and maintain. As a consequence, some improvising would become necessary in order to get information per type of copier to be used in discussion in small "type groups".

RESULTS

Resulting from this final stage in the development of ProMES are a three page individual feedback report and a two page region feedback report.

Moving averages. During the operationalization of the indicators it became clear that there would be large fluctuations in individual monthly indicator values (especially for percent repeat calls and percent return calls due to lack of car stock). The design team felt that feedback covering a larger period would be necessary to detect whether an indicator is really improving. A six month period was considered long enough to get a more stable measure. Using a moving average of six months also increases the reliability of the indicator values of the indicators measured by the supervisor of which it was doubtful whether the monthly sample would be large enough.

Percentage of maximum effectiveness. To be able to compare productivity of technicians, the monthly overall effectiveness score is expressed as a percentage of maximum possible effectiveness. This index is a measure of how well the technician is doing relative to how well he could be doing. This measure, in addition to the total effectiveness score had become necessary, because the possibility existed that data on some indicators were not available in a certain month. For example, a technician may not have had the opportunity to claim spare parts. In this case the indicator 'percent claims' does not apply to him/her and therefore neither an indicator value nor an effectiveness score can be obtained.

INDIVIDUAL FEEDBACK REPORT

The first page of the report showed how the total effectiveness score was composed of the effectiveness scores of products and indicators, for the last month and for the last half year.

The type-dependent indicators were described on the second page of the report. The monthly indicator values for the type-dependent indicators are calculated as follows. The total effectiveness score for an indicator was the weighted average of the effectiveness scores calculated for the different types of copiers. The weight was the number of calls the technician performs on each type, indicating the relative importance of each type of copier in that particular month. Only indicator values based on 5 or more calls on a type of copier are included. Suppose a technician performed 25 calls on type X with an effectiveness score of +20 on percent repeat, and 50 calls on type Y with an effectiveness score of -10 on percent repeat. In this case the overall effectiveness score on percent repeat would be 0.

As explained, some of the MCBC results originated from previous months, but were attributed to the month in which they can be calculated. Therefore, the calls used as weights in calculating the effectiveness score on MCBC were not the same ones as the ones used as weights for the other type-dependent indicators.

The moving average effectiveness scores were not directly linked to the moving average indicator values, because the contingencies were based on monthly indicator values. It would be incorrect to use these contingencies to calculate the effectiveness scores of indicator values averaged across a 6 month period, which have a smaller variance.

The composition of the effectiveness scores on the type-independent indicators was shown on the third page of the feedback report.

The region feedback report was an aggregate of the feedback reports of all individual technicians in the region. All technicians contributed equally to the region report. The region effectiveness score on each indicator was the average of the individual effectiveness scores on that indicator (except, of course, for the group level indicators). The structure of the region report was identical to pages 1 and 3 of the individual report.

PROGRAMMING OF FEEDBACK REPORTS

The technical implementation of the system was performed by a programmer from the data processing department, based on detailed specifications drawn up by the facilitators. This included programming of feedback reports and input and output facilities and took about 6 weeks. The facilitators entered the contingencies into the system. The result of the technical implementation is a program that generates (per month, per region) a set of individual feedback reports and a region feedback report.

FEEDBACK MEETINGS

PILOT MEETINGS

At the start of the implementation, the idea was to incorporate the feedback meetings into the monthly region meetings that were conducted by the supervisor.

However, it was decided to hold the first feedback meetings separately, because they would take up considerable time (about 2.5 hours). They were conducted in part by the facilitators, in part by the supervisor. The first 5 feedback meetings will be discussed. Preparation of feedback meetings. Each month, data on the individual indicators and group indicators were collected, and feedback reports were generated for all technicians in both regions. The feedback meetings took place between the 4th and 9th work-day of each month, to ensure timeliness of the feedback. It was decided to distribute the feedback reports during the feedback meetings. After each feedback meeting a memo was sent to all participants, confirming decisions and points for attention.

PROCEDURE DURING FEEDBACK MEETINGS

Region level indicators.

The supervisor's observations during the past month pertaining to "Compliance with maintenance procedure", "History Card" and "Ambassadorship" were discussed with the entire group. The supervisor explained how the effectiveness score on each indicator came about -using the contingencies- and compared it with effectiveness scores of previous months. He illustrated his evaluation with (anonymous) examples of positive and negative incidents he witnessed in the past month, some of which were shown on transparencies. Because of the routine nature of most elements of which the indicators consist, discussion was usually limited to directing attention to one or more areas that might be improved and getting group commitment on improving them.

Type independent individual indicators

In the initial discussions on the indicators "percent capacity used" and "percent claims (value)", it became evident that there were major problems with the accuracy of the data.

In the case of "percent capacity used", the facilitators tried to solve the problem by introducing the system to the planners in the central planning department who are responsible for entering data on the technicians' activities into the information system. As a result, the planners decided to check on missing data at the end of each month. Meanwhile, a number of technicians kept a written log on all time spent in order to check on the ProMES data. This resulted in accurate data by the fifth feedback meeting. However, even after such efforts, this indicator still remains very susceptible to errors, which can be very detrimental at the individual level.

"Percent claims (value)" turned out to be a hopeless case, primarily due to the complexities of the claiming procedure used by the department handling the financial aspects of claiming, which prevented accurate feedback to individual technicians. In all, these problems took up a lot of time during feedback meetings and thus left little time for

discussing the type-dependent indicators, which are considered the most important part of the system.

Type-dependent individual indicators

Individual interpretation of reports. During the third feedback meeting, the technicians were asked to individually interpret their own feedback reports using the contingencies and a description of the calculations involved. A number of questions was asked:

- Which aspects of the feedback report are not clear yet ?
- What does the feedback report tell you something about your strengths and weaknesses and about the strategies you follow ?
- Are you satisfied with your effectiveness score ? If not, what strategies could you try out to improve your effectiveness ?

According to expectations, the lack of participation of the technicians in the operationalization of the indicators and the complexity of this part of the report lead to a number of questions that could be answered by the facilitators, either immediately or after checking with the programmer.

There was, however, one exception: percent return calls due to lack of parts in car stock. The objective of this indicator -measuring to what extent a technician avoids return calls by maintaining a sufficient level of car stock- could not be achieved. Technicians are allowed to order car stock parts together with non-car stock parts. It cannot be determined whether the car stock part was ordered to finish the call or as a normal resupply of the car stock, thereby preventing accurate measurement a technician's supply control.

Visual representation of effectiveness scores. Although most technicians agreed that the feedback reports were useful tools to monitor and improve performance, the reports were considered complex and difficult to interpret, especially for the type-dependent indicators. For that reason, visual representation of effectiveness scores was added to the feedback reports. This consisted of line graphs representing individual and region effectiveness score over time for the type-dependent indicators. Using these graphs, the technicians could see at a single glance 'where they stand', and whether they have succeeded in improving both in absolute terms (compared to zero point) and in relative terms (compared to the region as a whole). Unfortunately, it was not possible to have the computer system generate the graphs. Therefore, the technicians had to fill in these graphs themselves. Although some technicians found this to be useful, others had serious reservations about this 'clerical' exercise. Although these graphs were just very simple, a lot of technicians for the first time got an idea about which -often implicit- strategies they had been using. For example, some technicians appear to have a 'high quality-high cost' strategy, whereas others appear to have a 'low quality-low cost' strategy. Because quality and cost are about equally important in the system, both strategies could in principle be

equally effective. By comparing their graphs with colleagues who have approximately the same models, strengths and weaknesses of individual technicians become apparent.

Discussion in small groups. The objective of small group discussion was learning about causes of performance differences between technicians by discussing individual performance data of those who work on the same types of copier. The focus should be on how to learn from each other and how to work more effectively and efficiently by using better strategies. Only during the fifth feedback meeting, was this type of discussion attempted. This was only partially successful. Some technicians did not like to talk about their individual performance. Yet, this had become necessary because of the lack of information per type of copier: technicians had to exchange a lot of information to get to know the extent of individual differences. Because of the limited experience with these discussions, no definite conclusions can be drawn.

In summary, the first five feedback meetings were partially successful. The part of the meetings that focussed on the group indicators progressed satisfactory. Yet, the recurring discussions on the validity of some of the indicators and lack of insight into the feedback reports hindered effective discussions on how to improve productivity on the individual indicators. Effects of the feedback meetings on productivity will be discussed later.

GENERAL INFORMATION

In this section, data are presented on time spent and costs involved in developing and implementing the system in East and Utrecht.

It took 16 months to develop the complete ProMES system, from June 1989 until September 1990. In both regions, there were 13 development meetings, including two review and approval meetings with management and one meeting to reach consensus between the regions. Developing the indicators turned out to be the most time-consuming activity, taking 2.5 and 3.5 meetings in East and Utrecht respectively. The development of contingencies took 3 meetings in East and 2 meetings in Utrecht. Latter numbers would have been much higher, if the scaling techniques had not been used. As is usually the case, developing products did not take up a lot of time (1 meeting in East and 1.5 meetings in Utrecht). There were review and approval meetings with both regions about the products and indicators, and about the contingencies. Before the second review and approval meeting, there was one meeting in both regions to achieve consensus on the system, and one inter-region meeting prepare for the second review and approval meeting with management. In the last development meeting, the design team developed the feedback report and discussed ways of handling the feedback meeting. In both regions, 33 hours were spent on the development meetings.

Some costs involved in the development meetings were:

- costs in connection with meetings (meeting rooms/aids/dinner): approximately \$ 16,000;
- payment of technicians and supervisors : approximately \$ 13,000.

The first feedback meetings were conducted in november 1990. Costs in connection with payment of technicians and rooms, aids and dinner were about \$ 1,000 per feedback meeting.

RESULTS

DESIGN

Effects of the system are evaluated using a time series design with 2 experimental units (East and Utrecht) and 9 control units (the nine other regions). In both experimental regions, there was only one month of baseline data (October 1990). In the experimental regions the system was implemented by means of monthly feedback meetings, during which there was feedback, but no specific goal-setting. In the nine control regions, data on all indicators were collected, but no information was fed back. The supervisors of these control regions were familiar with the PromES system. Their task was collecting the data on the group level indicators in order to get complete control group data. They received explicit instructions not to feed back any more information on these indicators than they have done in the past.

EVALUATION OF EFFECTS

The period under consideration is October 1990 until May 1991. Feedback meetings in East and Utrecht took place from October until March. The meetings in April and May were used to evaluate the system. The latter will be discussed later on, as well as results from june 1991 to april 1992.

Group level indicators (figure 2, left side, page 22)

In the experimental regions (East and Utrecht) the total effectiveness score on the group level indicators improved remarkably after the first feedback meeting (from -33 to 25). In the following months, the effectiveness score dropped a little and stabilized around 15.

The effectiveness of East and Utrecht during feedback is significantly higher than the effectiveness of the control regions. The small increase in the control regions may have been caused by the fact that collecting data on the group indicators supplied the supervisors with useful information.

GROUP INDICATORS

Maintenance Proc./History Card/Behavior

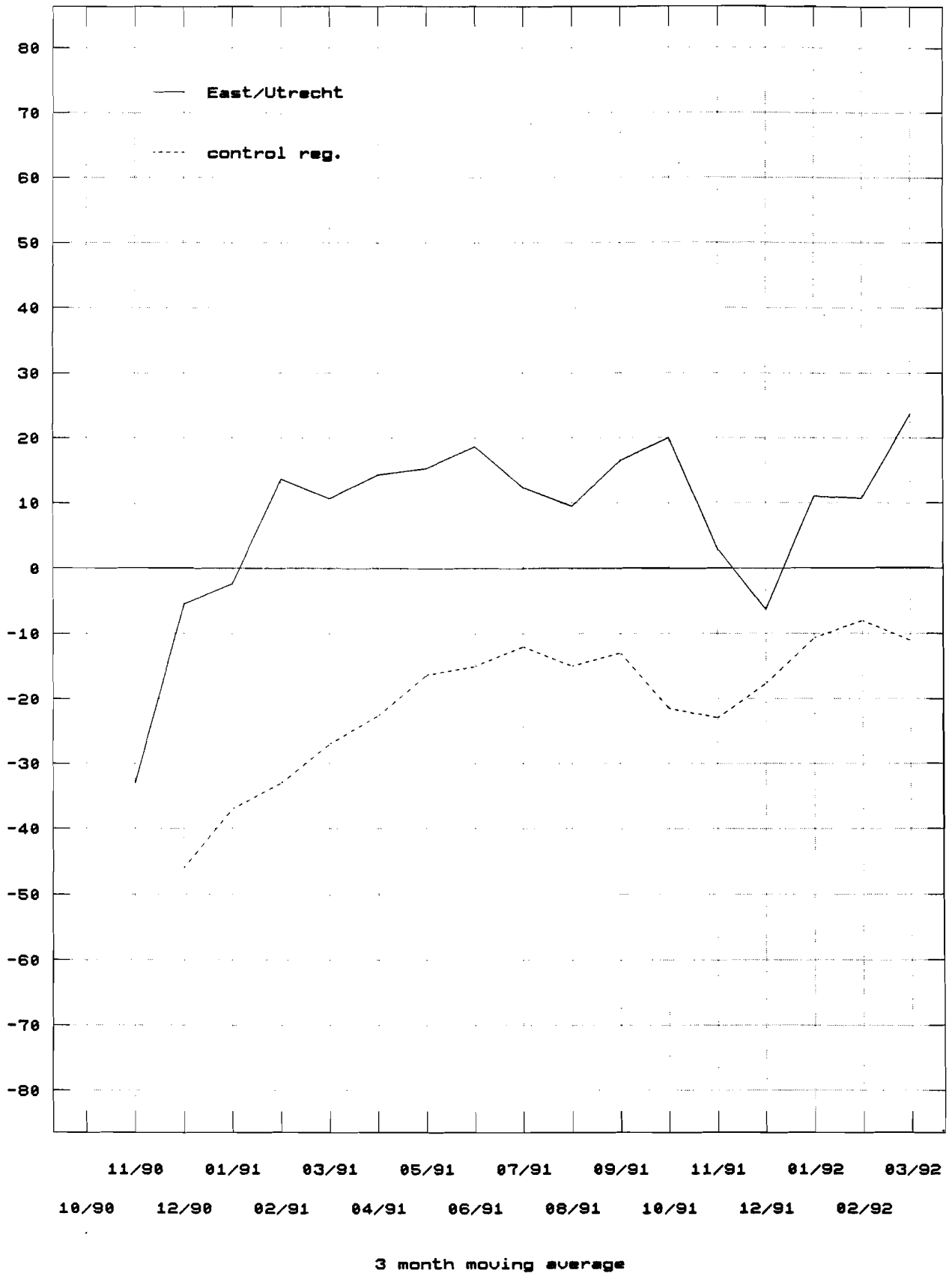


Figure 2: Effects of ProMES: group indicators

It must be noted that the feedback of information on group level indicators has had an immediate positive effect on productivity in Utrecht and a smaller and delayed effect in East. This was to be expected, since Utrecht's supervisor listed points for attention in a tell and sell way, which is considered effective, whereas the supervisor from East seemed less able to communicate the results clearly, especially during the first few feedback meetings.

Individual level indicators

Type-dependent indicators¹. The experimental regions show a slight increase in productivity (about 10 points) compared to the control regions. See figure 3, left side (page 24). The absence of large effects might in part be explained by the fact that relatively little attention had been paid to these indicators during the feedback meetings. Intensive group discussion, which was considered necessary to arrive at better strategies to improve productivity, hardly took place.

Type independent indicators. Due to difficulties in obtaining valid data on percent capacity used and percent claims, quantitative effects cannot be demonstrated. However, there is some qualitative evidence of some improvement in East and Utrecht. According to the planning department, technicians from East and Utrecht are concerned about attaining 100 percent use of capacity: they ask their planner for an extra call if they feel they might not be able to 'make their hours'. Also, during the feedback meetings it became clear that the technicians have been paying special attention to claims in order to check the validity of data that is fed back.

DISCUSSION

Although there were improvements made in group level effectiveness, both management and facilitators were disappointed about the overall effects of the feedback provided by the ProMES on the performance of the technicians who had developed the system. The following steps were taken to solve this problem.

First, the basic issue of potential performance improvement was settled: technicians were of the opinion that, notwithstanding the effects of performance improving measures taken at earlier times by the management, there still was room for additional improvement. The management, when asked so informally, was of the same opinion.

¹Excluding percent return calls parts car stock, because of difficulties in obtaining valid measures. Evaluation of effects includes those individual indicators, which were considered valid from the start of the feedback period.

INDIVIDUAL INDICATORS

MCBC/repeat/labor time/parts cost

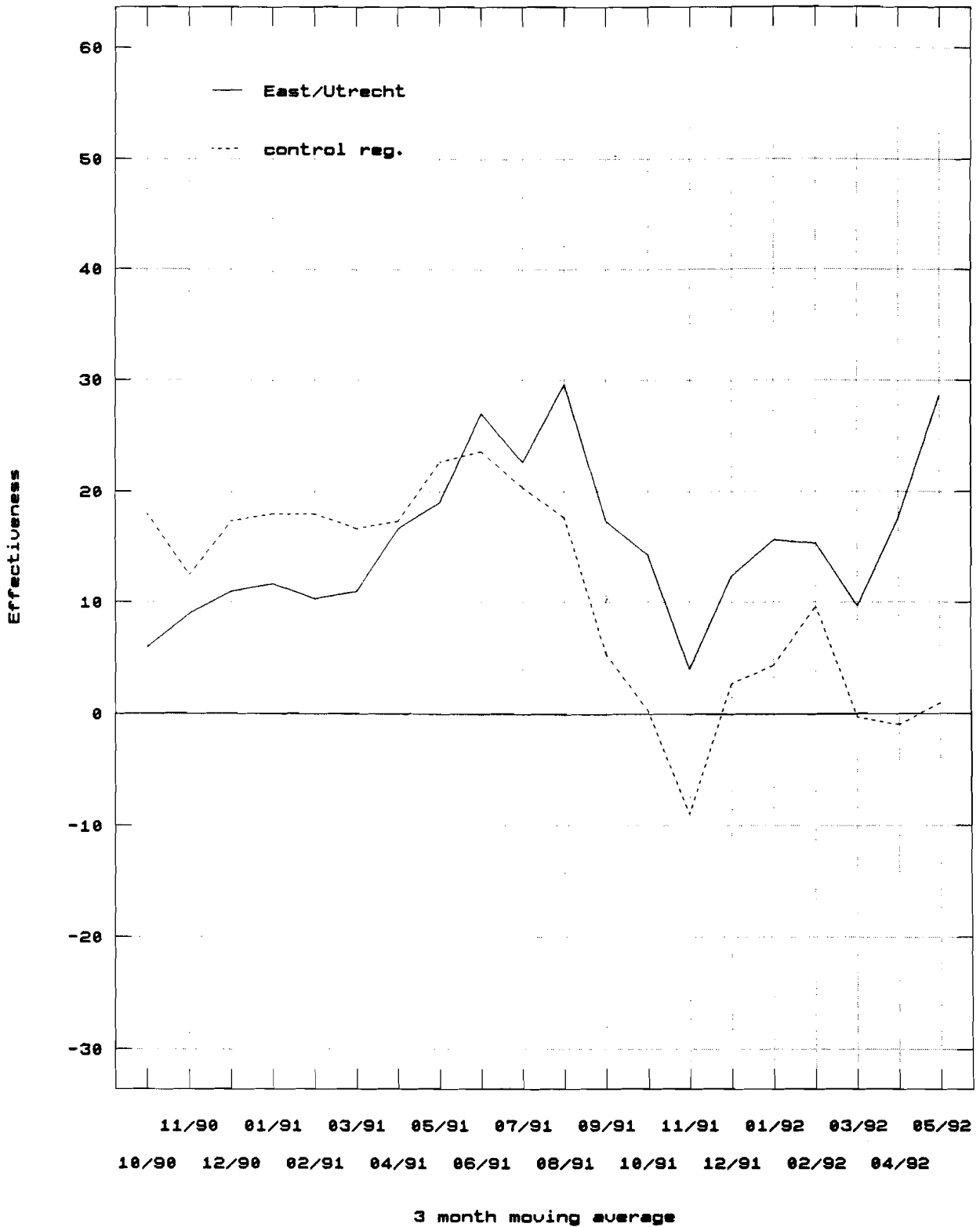


Figure 3: Effects of PromES: individual indicators

Although they could not give indications as to the degree of potential performance increase, they could give anecdotal evidence indicating that at least some efficiency improvement had to be possible.

As a next step, a list of possible causes for the absence of substantial performance improvement on the individual indicators was made with the help of supervisors and technicians. The facilitators added some factors to the list on the basis of theoretical considerations. The factors were discussed in each design team and rated according to perceived importance.

After that, the design teams together with the facilitators developed the following list of solutions, which corresponded to the main perceived causes of the problem. First, a clear management statement on ProMES was asked for. Technicians felt kind of left alone, working during a very long period on developing ProMES without getting regular signals from the management that this was still appreciated and in conformance with management policies and intentions. "Is this still serious, or do they have other interests and priorities right now? If so, why should we bother?" Secondly, technicians asked for the removal of invalid indicators from the feedback reports. Lots of discussion time during feedback meetings was spent on validity issues. Many times, technicians with or without the help of the facilitators, successfully demonstrated that data in the feedback reports were incorrect. This resulted in feelings of distrust which did not stimulate performance improvement initiatives. Besides, not much time was left for discussing such initiatives. Thirdly, feedback reports were considered too complex. Technicians got overall ProMES effectiveness score's on each indicator and per indicator, for each type of photo copier separately, both indicator score's and ProMES effectiveness score's. To be able to discern trends, graphic presentation of this information in fact was a must, but it was not available. Therefore, it was proposed to present information on the valid indicators both graphically and tabular. The fourth solution also considered the feedback report. Up till now, performances on the indicators measured samplewise at the level of the region were included in the individual feedback reports of each technician, thereby suggesting that the scores represented the individual performance of the technician during the period concerned. Technicians refused to be held personally responsible for the performance of their colleagues and asked for removal of these indicators from their individual reports. The solution proposed was to present a feedback report containing overall region's scores on all valid indicators (including those measured only samplewise) to the whole group, and individual feedback reports containing only information on the valid individual indicators, to each individual technician. These latter reports could be discussed in bilateral meetings of an individual technician with his supervisor, whereas the region report could be discussed in a feedback meeting of all the technicians of the region with their supervisor. A fifth cause played a role, but did not come explicitly to the surface until the management reacted to the first four solutions proposed by the design teams. It had to do with the issue of "what is in it for us, if we succeed in improving our

performance?". In fact, opinions on this point were mixed, ranging from "you are already paid for doing a good job, so why should you get more for eventually really doing a good job" till "if we improve, the company earns a lot more money, why shouldn't we get our share; besides, it stimulates to have an outlook on a reward".

The proposed solutions were brought to the attention of the management and management agreed to arrange a meeting at which they would make a statement and react to the concrete suggestions of the design teams. The management statement at these meetings came as a surprise. In short, they took the position that, although ProMES had not brought the improvements expected, it was still very worthwhile to have it and use it for performance appraisal purposes. This statement caused a lot of confusion, given the guarantee at the start of the project to not do such a thing unless all parties concerned would agree. In a long discussion, several of the ingredients presented before were combined into a plan acceptable for all. To begin with, although at first outraged the technicians soon understood that the management proposal fitted in very nicely with their fifth problem, the absence of a link between performance improvement and reward. It was agreed that their annual bonus would be linked to their performance as measured by ProMES. This would be done by making ProMES an integral part of the performance appraisal system, by means of which the amount of annual bonus is determined. The exact procedure would be worked out and presented to the technicians for approval. A precondition was that only valid information would be used, so the discussion continued on the other solutions proposed by the technicians. In fact, all suggestions made were accepted by the management. In short, agreement was reached on the content of the individual feedback reports: only the indicators the validity of which was accepted were included. These are 'Mean Copies Between Calls', 'percentage repeat calls', 'parts cost per call', 'labor time per call'. That is to say that only the two more important products, quality and cost, are included in the individual reports. The region report will contain information on the overall score of the region on those four individual indicators, and in addition to that the score's on 'compliance with preventive maintenance', 'accuracy of history card' and 'correctness of behavior', the three indicators measured samplewise by the supervisor. The request for graphic information presentation was granted too. Finally, it was decided to have regular meetings with the management in the future to evaluate the progress on the ProMES project and to discuss mutual affairs.

Figure 3 (page 24) shows the performance of the two trial regions and that of the other regions on the four individual indicators during the period in which feedback was provided, the period during which the discussions on lack of improvement took place, and the period after the agreement with management was reached. It can be seen from the graph that whereas in the first period the trial regions are below the other ones, their performance is superior in the period after the final agreement with management was reached.

So, after removal of the main causes of dissatisfaction with the system, at least some improvement occurred on the individual indicators. In addition to that, the performance increase on the group indicators was maintained. A third positive result to be mentioned was the high opinion the management declared to have of the measurement part of the ProMES system: they were convinced that the complexities of the job of the service technicians were accounted for extremely well by the system and that, after the revisions described before, the system was valid. They were particularly enthusiastic about the way an until now unsolvable problem, viz. the incomparability of performances of technicians on different types of photocopiers, was taken care of by introducing a separate set of contingencies for each type of copier. Not only the management, but all parties concerned were of the opinion that the measurement system offered by ProMES was a sound one, and far better than anything else used in the past. This was the main reason for management and technicians to agree to use the technicians' individual ProMES scores as input for the performance appraisal system, and as a basis for determining the amount of annual bonuses.

Although many positive results have been accomplished in this case, we are still left with the question why there is far less performance improvement than expected. So, which are the factors inhibiting success? In discussing potential inhibiting factors, we would like to use the control loop model as a metaphor. The essential elements of a control loop are norms or goals, and feedback. ProMES can be considered a procedure to develop and implement accepted control loops. According to such a line of reasoning, ProMES can only be effective to the degree that the feedback provided by the system that is developed, as well as the goals implied by the system, are accepted by those who are supposed to control their own performance by means of the system. Discrepancies between accepted goals and accepted feedback would result in attempts to reduce these discrepancies by investing effort in strategies leading to high performance.

The above mentioned model helps us to explain why the effect sizes in this case were small compared to what has been found in other cases. In other words, it helps us to organize what we have learned in this case. For several reasons, feedback acceptance might have been low. As discussed above, the feedback reports were rather complex. They provided indicator data for each of the types of photo copiers serviced by a technician, and combined these data through the contingencies that belonged to each type into overall effectiveness scores for each indicator and each type of copier. In addition to that, six months moving averages were provided. Some information was given as raw indicator data, other in the form of effectiveness scores. For most indicators, information referred to the prior month. The MCBC indicator, however, referred to an undefined period in the past. For these reasons, most technicians had difficulty understanding the system, notwithstanding the fact that they had assisted in developing it themselves. It will, no doubt, be difficult to accept and use feedback that one does not understand completely.

Acceptance will also have been lowered by recurring discussions on the validity of some indicators.

For accepted discrepancies between feedback and goals to occur, in addition to accepted feedback accepted goals are a prerequisite. It is doubtful whether there were accepted goals in this case. The main reason for this stems from the way in which the contingencies were developed. In estimating maxima, zero points and minima, the performance data from all (about 300) of the company's technicians working all over the Netherlands were used. By implication, many technicians belonged in the lower half of this distribution, and many in the upper half. The former group will receive mostly negative effectiveness scores, the latter group positive ones. For this reason, a lot of technicians will have difficulty in accepting the maxima, zero points, and minima as relevant for themselves as individuals. As a consequence, they will have difficulty in defining relevant performance ranges for themselves. In other words, instead of helping them to establish specific, difficult but attainable goals, the contingencies might very well confuse technicians about the goals to be strived for.

Accepted discrepancies between feedback and goals should result in the decision to invest effort in a strategy that leads to a maximum performance increase in the next period. In this case, however, to decide for such a strategy is a complicated affair, because the contingencies of all the types of copier the technician is servicing have to be taken into account. In fact, there are several options. The technician could choose to improve his/her performance on the type on which improvement on all the indicators would result in the largest increase of his/her overall effectiveness score. The technician could also decide to improve his/her performance on the indicator on which an overall improvement on all the models would result in the largest increase of his/her overall effectiveness score. Of course he/she could decide to use some combination of the two options mentioned. Because of the large number of variables involved, it is very difficult to determine the best strategy to follow (it is even hard to determine a good or acceptable one). In fact, the feedback report should help the technician by providing him/her with predictions on how his/her overall effectiveness score would behave in case one or the other strategy would be followed. However, the feedback report does not contain that facility yet, because we lack a good optimization model.

Taken together, the complexity of the ProMES system developed in this case seems to be responsible to a large degree for the absence of large performance increases.

References

- Edwards A.L., *Techniques of Attitude Scale Construction*, New York: Appleton-Century-Crofts, 1957.
- Matsui T., Kakuyama T. & M. Onglatco (1987), *Effects of Goals and Feedback on Performance in Groups*, *Journal of Applied Psychology*, 72 (3), 407-415.
- Mitchell T.R. & Silver W.S. (1990), *Individual and Group Goals When Workers Are Interdependent: Effects on Task Strategies and Performance*, *Journal of Applied Psychology*, 75 (2), 185-193.
- Pritchard R.D. (1990), *Measuring and Improving Organizational Productivity: A Practical Guide*, New York: Praeger.