THE EFFECTS OF A VISUAL FEEDBACK DISPLAY ON SPIROMETRIC TEST PERFORMANCE

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

by

Charles Marlin Brown

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Psychology

Georgia Institute of Technology

September, 1977

THE EFFECTS OF A VISUAL FEEDBACK DISPLAY

ON SPIROMETRIC TEST PERFORMANCE

Approved:

Randall M. Chambers, Chairman

Charles V. Riche

James M. Bradford

Date approved by Chairman: 9/21/77

ACKNOWLEDGMENTS

I would like to express special thanks to all members of my thesis committee. Dr. Chambers has provided immeasurable guidance and support in this and other of my academic efforts. Dr. Riche's help in design and statistical analysis as well as his tireless work in helping to finalize the write-up was invaluable. Without Dr. Bradford's support, research on this topic would have been impossible. He made the sophisticated research equipment, funding, and his own programming skill available without hesitation.

Thanks are also due to the National Institute for Occupational Safety and Health (NIOSH), who funded the equipment used, and to Dr. Alan Palmer of NIOSH, who provided a demonstration of his own spirometric feedback system and ideas which aided in the development of mine.

I am also grateful to Mrs. Joyce Williams for her help in typing and preparing the manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This study tested the effectiveness of a visual feedback display in improving subject's performance in spirometric testing. The test was considered as a motor response and the principles of psychological feedback were applied to aid in the acquisition of this response.

The display provided feedback concerning the adequacy of expiratory effort with regard to (1) volume of air exhaled, (2) maximum flow rate achieved, and (3) time of sustained effort. Four groups of subjects received either (1) no verbal or visual display feedback, (2) verbal feedback from the technician according to standard spirometric procedure, (3) visual feedback from the display, or (4) both visual display feedback and verbal feedback from the technician.

The results showed that the use of the display yielded better mean performance on vital capacity and at least as good as standard spirometry on forced expiratory volume in one second, peak flow rate, the time of the vital capacity, and the flow rate at 50 percent of the vital capacity. A marked reduction in variability also resulted from use of the visual display. For the visual display group the within-cell variability on percentage of predicted normal vital capacity was less than one-half the variability of the standard verbal feedback group. On percent of predicted forced expiratory volume in one second both the display group and the display plus verbal feedback group showed less than one-fourth the variance of the verbal feedback group.

These results suggest that use of a display like this one as a

standard part of a pulmonary function testing system would increase the quality and reduce the variability of the test results. This could lead to more accurate medical diagnoses based on the test and, in experimental studies, a greater sensitivity of the test to treatment effects.

CHAPTER I

INTRODUCTION

Engineering psychology is a field that has been characterized by work at the interface of other disciplines and areas of research since its earliest days. Useful and important applications of the findings and methods of experimental psychology in industrial design, industrial engineering, and systems design and management are well known examples of the fruitfulness of psychology's work at the interface of other fields.

Recently, as pointed out by Alluisi and Morgan (1976) in their review of the present status of engineering psychology, much interest and research has emerged in the application of psychological research methods to problems in the health field, particularly occupational safety and health. This is evidenced by a large number of recent publications, particularly from the Behavioral and Motivational Factors Branch of the National Institute for Occupational Safety and Health (NIOSH), reporting important findings from behavioral research on job hazards, job demands, safety practices, occupational exposure to toxic substances, and the health, safety, and performance of the worker (Cohen, Smith, and Cohen, 1975; Repko, Morgan, and Nicholson, 1975; Prather, Crisera, and Fidell, 1975; Sleight and Cook, 1974; Caplan, Cobb, French, Harrison, and Pinneau, 1975; Xintaras, Johnson, and de Groot, 1974).

The study reported in this paper extends the application of psychological research in the field of medicine to improving the quality and reliability of a specific dependent variable in the field of medicine, the forced expiratory volume spirometry test. Specifically the use of a visual display to provide knowledge of results to the patient or subject about the adequacy of his effort or performance in this test should considerably reduce many of the reliability problems of spirometric testing.

This study is to test the usefulness of psychological principles of augmented or information feedback in increasing the quality, reliability and ease of administration of respiratory function testing by spirometry or plethysmography, which are commonly used in mass screening for occupationally induced respiratory disease. The most important breathing maneuver in these tests requires a maximum forced exhalation by the subject after a full inspiration. To be able to interpret the results of this test accurately, the physician must feel certain that the subject has understood the instructions, knows what he is supposed to do, and is motivated to provide a truely maximum effort in the forced expiration. Otherwise, the test is unreliable and possibly misleading. These problems of effort dependency and test reliability will be discussed in detail in the next section.

This study investigates the possibility that a visual feedback display cued to the subject's own performance during the spirometric test can facilitate his comprehension of the test instructions, motivate him to provide a maximum effort, simplify the test protocol for the technician, and help to minimize effects due to differences among technicians or within the same technician over time by providing a more constant, programmed test protocol that is less dependent upon the technician.

CHAPTER II

QUALITY CONTROL PROBLEMS IN SPIROMETRIC TESTING

The Test

The forced expiration test requires the subject to inhale maximally, then exhale with as much force and as rapidly as possible, continuing the expiration until he can move no more air (West, 1974; Ruppel, 1975). The total volume of air the subject is able to exhale is called the forced vital capacity (FVC) or simply vital capacity (VC). The second basic volume parameter is the volume the subject can exhale in the first second of the maximum effort expiration, called the forced expiratory volume in one second ($FEV_{1.0}$).

Electronic spirometers include transducers to measure instantaneous flow rates as well as volumes. Measurement of flow rate patterns at various cumulative volumes or times in the test can give much additional information to the physician about disease, constrictions and obstructions in the airways of the lungs. Results of forced expiratory tests are often displayed as flow-volume curves, which are plots of the instantaneous flow rate at each volume of expired air for the course of the test. A typical flow-volume curve for a normal subject is shown in Figure 1. The shape of the flow-volume curve provides valuable information to the physician trained in reading these curves. Some respiratory diseases cause characteristic deviations from the normal shape of this curve (West, 1974).

Two important respiratory parameters are calculated on the basis

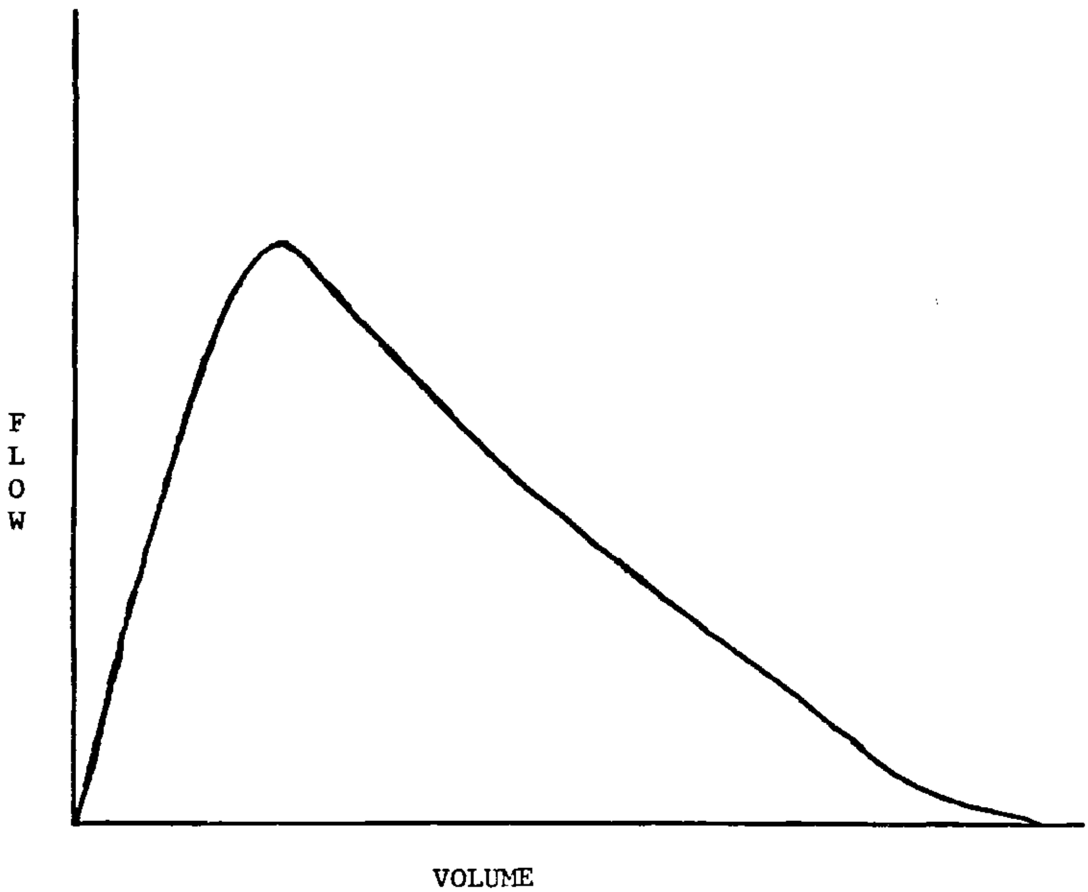

Figure 1. A Typical Flow-Volume Curve of a Forced Expiration by a Healthy Subject

of flow rates: (1) the highest instantaneous flow rate obtained, called the peak flow (PF), and (2) the instantaneous flow rate at 50 percent of the vital capacity ($\dot{V}_{50}$). The time of the vital capacity (VCT) is the time in seconds from the start until the end of a maximal expiration. VCT can serve as an indication of the extent to which the subject sustains effort towards a full expiration (Ruppel, 1975; West, 1974). Table 1 summarizes these respiratory terms.

A behavioral analysis of the properly performed forced expiration shows the following response sequence: (1) The subject performs a slow maximal inspiration, (2) without hesitating after reaching his full inspiratory capacity, he starts the expiration by exerting maximal force thrust with his diaphragm and chest muscles to force air out of his lungs as rapidly as possible, (3) he sustains maximal effort through the decelerating air flow of midexpiration, and (4) he sustains effort through the slow, asynchronous emptying of his lungs until he can move no more air. The adequacy of the subject's performance at each of these stages of the expiratory response, and thus the meaningfulness of the test results, can be affected by the subject's experience with the test, his comprehension of the instructions, his willingness to exert a maximal effort, and the ability of the technician to motivate full cooperation and effort.

Practice Effects

Many studies have shown that the reliability of forced expiratory spirometry tests is usually quite a problem. Discher (1970) reported that when subjects were retested two to six months after initial testing, test-retest reliabilities of .92 for $FEV_{1.0}$ and only .79 for FVC

Table 1.

Definitions of Respiratory Terms Used

Abbreviation	Term	Definition
1. VC FVC	vital capacity forced vital capacity	maximum amount of air that can be exhaled
2. $FEV_{1.0}$	forced expiratory volume in 1 second	volume exhaled in first second of of forced expiration
3. PF or $\dot{V}_{max}$	peak flow rate	highest instantaneous flow rate achieved in a forced expiration
4. $\dot{V}_{50}$	instantaneous flow rate at 50% of vital capacity	
5. VCT	time of vital capacity	time in seconds from start to end of forced expiration

were found with a systematic tendency towards increasing volumes for both $FEV_{1.0}$ (79.1 ml.) and FVC (148.2 ml.) on retest. This general trend towards higher volumes is probably attributable to experience and practice. Many subjects whose initial tests fell below the cutoff values used to classify them as "positives," suggesting the presence of respiratory disease, were classified as "negatives," or not diseased, on retest. Discher concluded that the test is only moderately repeatable and further efforts to improve its reliability are needed.

Another study of the reliability of the spirometry test by Discher, Massey, and Otoupalik (1970) reported increases on retest not only in the $FEV_{1.0}$ and FVC volume parameters, but also significant increases in the maximum midexpiratory flow, maximum expiratory flow rate, and peak flow parameters. The authors suggest that the retest measures are more representative of the subjects physiological breathing capacities and that previous experience with the test accounts for the increased performance on retest. The change in performance on retest caused 37 percent of the subjects classified as positive on the first test to shift to negative on retest. Subjects were categorized as positive if their FVC or $FEV_{1.0}$ volumes fell more than 1.645 standard deviations below the predicted normal value based on regression equations using age, height, and sex to predict respiratory parameters (Kory, 1966).

The Discher et al (1970) study suggests three related factors contributing to this lack of reliability: (1) the lack of full comprehension of test instructions by subjects, (2) the lack of motivation to exert a maximal effort, and (3) the effects of practice in learning the motor responses required for maximal inhalation and maximal forced exhalation. All of these problems can probably be addressed more effectively from a psychological point of view than from a medical one and the use of a visual feedback display seems a fruitful approach to inform and motivate the subject and to facilitate his learning of the responses required of him.

Effort Dependency

Dayman (1967) has studied the problem of effort dependency in the forced expiration test. He described three phases of the flow-volume curve. Phase I flow rates are highly dependent on the amount of effort the subject supplies in the initial blast of the forced expiration. This phase starts at the beginning of the expiratory maneuver and lasts until about 35 percent of the vital capacity has been exhaled. Phase II, which lasts from about 25 percent to 75 percent of the vital capacity is characterized by constant deceleration and is relatively independent of the amount of effort exerted by the subject. Phase III is the slow asynchronous emptying of the last quarter of the vital capacity. This phase is highly dependent on the sustained effort of the subject. Premature termination of effort in this stage can cause underestimation of the subject's vital capacity.

Discher and Palmer (1972) specify several of the effort related problems in spirometry. (1) The subject might not reach a full inspiration before starting the forced exhalation. (2) He might not exert a maximum effort in the initial thrust. (3) He might hesitate or inspire in mid-expiration. (4) He might terminate expiration before he has exhaled his full vital capacity. (5) He might produce artifacts by a loose seal on the mouthpiece or by pursing his lips or tongue, causing an inaccurate flow measurement.

Technician Effects

The strong dependence in spirometric testing on subject motivation, cooperation, and comprehension of instructions makes the technician who administers the test an important factor in determining the quality and reliability of test results. The technician must act as "a bully, cheerleader, and psychologist as he strives to elicit a maximal response from the subject" (Palmer, Ayers, Abraham, and Wilbur, 1971). These are skills which the technician must acquire through experience. Large individual differences among different technicians in their ability to motivate and instruct subjects also affect the reliability of test results. Performance of the same technician will also vary with time as he becomes fatigues or bored with repeated testing, frustrated by uncooperative subjects or hoarse from the loud verbal exhortation required to motivate maximal effort.

Discher, Massey, and Hallett (1969) found that with experience, the technicians' ability to elicit satisfactory tests increased. For the first two days of testing the three technicians in this study showed an average unsatisfactory test rate of 24.5 percent while on days three through six their average performance improved to generate only 15.6 percent unsatisfactory tests. A marked increase in the percentage of subjects showing normal values on respiratory parameters was also associated with technician experience.

Palmer, et al (1971) report that for experienced technicians the sex of the technician can affect spirometry results. Two male and two female technicians tested 1015 male subjects. Male technicians elicited significantly higher early flow rates than did female technicians. This means that subjects put more effort into the initial blast after inhalation when the technician was a male. This study leaves unanswered the question of what effects the sex of the technician has on spirometric tests of female subjects.

These technician effects underscore the potential usefulness of a visual feedback display. If a display can effectively reduce the amount of technician intervention required in spirometry many of these inter-technician and intra-technician effects might be minimized. The test conditions would be less variable between technicians and over time. In addition, prolonged testing should cause less fatigue and strain on the voice of the technician if the feedback display can take over some of the load in motivating and instructing the subject.

The visual feedback display should reduce practice effects by facilitating the acquisition of the adequate response within the standard five spirometry trials. If higher values are shown by subjects with a feedback display than by those with no feedback, this means the measured value is closer to the subjects' physiological capacities. The feedback display might also aid the subjects in comprehending fully the test instructions.

The problem of effort dependency should also be reduced by the probable motivating and interest-catching effects of the feedback display. The subjects may view the test as a game in which they are competing for a "perfect score" of all lights lit.

CHAPTER III

FEEDBACK APPLICATIONS IN RESPIRATORY PHYSIOLOGY

Studies applying the principles of psychological feedback to problems in the field of respiratory physiology in general are quite rare and applications to spirometry in particular are even less frequent. The literature on biofeedback (Brown, 1975, offers a good bibliography) yields little of direct relevance because breathing is a response which is normally under voluntary control and biofeedback studies focus on voluntary control of autonomic, normally involuntary responses.

K. U. Smith and his colleagues (Henry, Smith, and Rosenstein, 1966; Henry, Junas, and Smith, 1967; Smith and Henry, 1967) have studied feedback in breath pressure control and control of ventilation rate for normal subjects and for emphysema patients. Their results show strong effects of visual feedback on breath control, but the emphasis in these studies has not been as much on the use of feedback to improve breath control as on the use of breath control as a means to study delayed feedback.

Block, Lagerson, Zohman, and Kelly (1969) have reported the successful application of feedback techniques to training patients in diaphragmatic breathing. For patients with chronic pulmonary diseases of several kinds, it is considered desirable for the patient to learn to change from a thoracic to a diaphragmatic mode of breathing. In this study patients were provided feedback from a red light and a buzzer to indicate when they were breathing incorrectly. The feedback device proved highly effective in training patients to breath with the diaphragm rather than the chest.

Feedback in Spirometry

Subjects receive some feedback, that given verbally the the technician, in all standard spirometric tests. The problem here is that the accuracy, information content, and instructional effectiveness of this feedback is variable from one technician to another and from one test to another, as discussed previously. The usefulness of augmented feedback from a visual display has been suggested by the few studies that have tested this approach.

Palmer, Ayers, Abraham, and Wilbur (1971) used a visual display composed of four lights to provide the subject feedback on his performance. Three of these lights were cued to light up when the subject reached 80 percent of his predicted normal value on peak flow rate, forced expiratory volume in one second ($FEV_{1.0}$), and forced vital capacity (FVC). The peak flow light required a strong effort on the initial thrust and the $FEV_{1.0}$ and FVC lights required a sustained expiratory effort to light them. The fourth light, and end of test light, encouraged maximal lung deflation by lighting up only if the subject maintained his expiratory effort for four seconds after the start of the forced expiration.

Subjects in the control group (no visual feedback display) were given standard spirometry instructions at the start of the test and verbal exhortation from the technician during the test to encourage a maximal effort. The experimental group was given similar instructions at the beginning of the test but no verbal encouragement during the test. They were told to watch the lights and to try to light all of them by exerting a maximal effort. They were not instructed concerning the differences among the four lights, only to try to light all of them.

The results shows (1) no differences due to the display in several volume measurements, including FVC and $FEV_{1.0}$ and (2) significant increases in several flow rates due to the feedback display: $FEF_{25\%-75\%}$, $FEF_{25\%-50\%}$, $FEF_{50\%-75\%}$, and flow rate at midexpiration.

Discher and Palmer (1972) describe a revised model of the feedback display described above, with more lights cued to fewer parameters. Seven lights were used, one end of test light cued to require a four second sustained expiration, and six lights all cued to successive fractions of the subjects predicted normal vital capacity. This article simply describes the system with no experimental tests of its effectiveness reported.

The Feedback Display

Some weaknesses in the two visual feedback display systems described above suggest that a better design may be more effective. The four light system provides all-or-none, qualitative, imprecise feedback information for each of the four parameters sampled, since only one light is cued to each of these parameters. The authors recognize this weakness by pointing out that a patient with respiratory disease might not be able to meet the 80 percent of normal criterion on any of these parameters and may thus receive no feedback at all (Palmer et al 1971).

A second problem with the four light display system is caused by not giving the subject instructions to discriminate what the lights cued to different parameters represent. The FVC and end-of-test lights require sustained effort of the subject while the peak flow light requires a maximally forceful initial thrust. These seem to be different dimensions of the forced expiratory response and the feedback should allow the subject to discriminate whether he is not blowing long enough or not blowing hard enough if he fails to light all the lights.

The problem of feedback not representing the dimensions of the response also applies to the seven light system. Here, feedback on the adequacy of the initial thrust is not given at all, since no flow parameters are represented in the display. The problem of supplying quantitative, more precise feedback is addressed, however, by cueing six of the seven lights to successive fractions of a normal FVC.

If the forced expiratory maneuver is viewed from a psychological point of view as a motor response which we want to help the subject learn as quickly as possible, some general principles of information feedback and knowledge of results can be applied to this specific situation.

A long established, fundamental psychological principle states that performance can be improved and acquisition of a response accelerated by providing knowledge of results and that the effectiveness of this feedback depends on the precision of the knowledge of results (Thorndike, 1927; Trowbridge and Cason, 1932; Elwell and Grindley, 1938; Macpherson, Dees, and Grindley, 1948; Bilodeau, Bilodeau, and Schumsky, 1959). This suggests the benefit of a visual display to aid the subject in learning to exert maximal effort in the forced expiratory response. It also indicates that the most effective feedback display should be one providing quantitative and precise information to him about the nature of any inadequacy in his performance.

In both the spirometry feedback studies reported here the subjects were simply instructed to try to light all the lights with no indication to them of what the lights meant or what they should do to light them. Although the lights were cued to different parameters requiring different response dimensions, the subjects were given no means of discriminating among the lights and thus no means of knowing why they failed to light all the lights. If the subject fails to exert sufficient force in the initial blast of the expiration, the peak flow light in the four light system will not ignite. If he starts the expiration after a submaximal inspiration, the VC light(s) will not be lighted.

With these three dimensions of the response in mind (maximal force in initial thrust, full inspiration before starting expiration, and sustained effort until a full VC is emptied), a display with discriminably different lights cued to peak flow, vital capacity, and time of expiration was designed, built, and programmed. The instructions given to the subject and labels under the display lights tell the subject (1) to exhale with greater force if he fails to ignite all peak flow lights on the preceding trial, (2) to inhale maximally if he doesn't light all VC lights, and (3) to sustain his effort longer if he doesn't light the timed end of test signal light.

Feedback on volume and flow rate adequacy was given by four lights for volume and four for flow, with each light cued to a graded percentage of the subject's predicted normal vital capacity or peak flow. The four lights were cued to ignite in sequence as the subject achieved

70, 80, 90 and 100 percent of his predicted normal peak flow and, similarly, the four volume lights were cued to increasing percentages of predicted normal vital capacity. The normal values were derived from regression equations using age, hgieht, and sex to predict respiratory parameters for healthy subjects (Leiner, Abramowitz, Small, Stenby, and Lewis, 1963; Morris, Koski, and Johnson, 1971). This graded, quantitative feedback will allow the subject to perceive his improvement as he approaches criterion performance over trials and should be more effective in eliciting maximal performance than the Palmer et al (1971) four light display, which allowed only one light for each respiratory parameter.

The display is shown in Figure 2. It contains four green lights in one row cued to successive percentages of normal peak flow, four blue lights in another row cued to percentages of normal vital capacity, and one red end-of-test light set to ignite if the subject maintains his expiratory effort for at least 0.5 seconds after his flow rate decreases to 0.45 liters/second.

To evaluate the usefulness of this display, four groups of subjects were tested under these conditions: (1) the no feedback group received neither verbal feedback from the technician nor visual feedback from the lights display, (2) the verbal feedback group received verbal coaching and encouragement from the technician according to standard spirometry procedures, (3) the lights feedback group received feedback only from the visual display, and (4) the combined feedback group received both verbal coaching from the technician and feedback from the lights display.

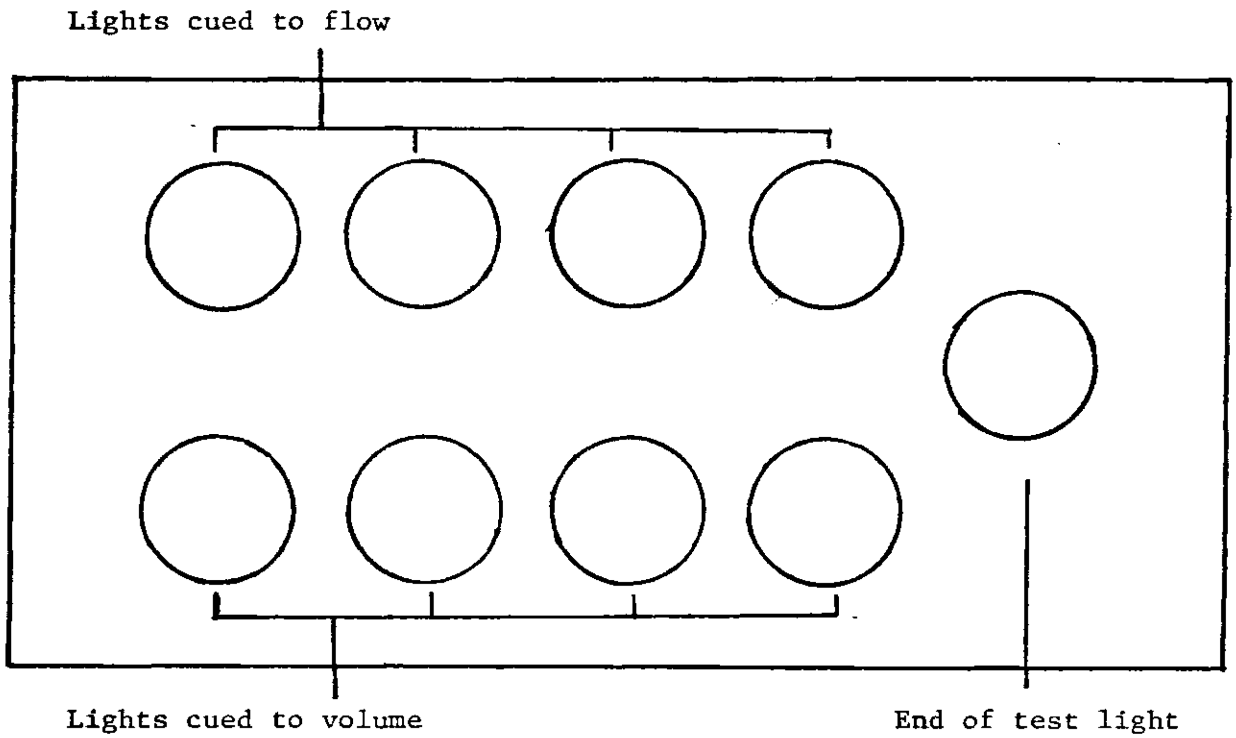

Figure 2. Visual Display Panel

Hypotheses

The following hypotheses were formulated: (1) The groups receiving visual feedback (lights and combined feedback groups) should perform better on several important respiratory parameters than either of the groups not receiving visual feedback (verbal and no feedback groups). (2) Considering the dependent variables separately, the groups receiving feedback from the lights display should perform better than the verbal and no feedback groups on percentage of predicted normal vital capacity (PVC), forced expiratory volume in one second (PFEV1), peak flow rate (PPF), and time of the vital capacity (VCT). The flow rate at 50 percent of the vital capacity should not be affected by feedback since the test is relatively effort independent at mid-expiration (Dayman, 1967). (3) By reducing technician intervention the visual feedback display should also reduce variability in PVC, PFEV1, and PPF.

CHAPTER IV

METHOD

Subjects

Sixty students from psychology classes served as subjects. Of these 19 were females and 41 were males. They received extra class credit for participating in the experiment. All reported having no chronic respiratory disease and all but four were nonsmokers. Consent documents explaining fully the nature and purpose of the test were read and signed by all subjects who were tested.

Apparatus

The respiratory testing equipment to be used in this study is installed in a mobile pulmonary function testing laboratory maintained by Emory University under a research grant from the National Institute for Occupational Safety and Health. The laboratory is totally self-contained in a recreational vehicle chassis. The testing equipment can be operated either from an external power source or by on-board electric generators. The on-board equipment used in the present study includes a body plethysmograph, which was used to measure volume and flow rates in the forced expiratory test, and a PDP 8-E minicomputer, which was used to program the testing sequence, monitor and record the data on magnetic tape, and control the presentation of feedback stimuli for the visual display. The computer calculated the predicted values for each subject on vital capacity and peak flow, using these values to set the criteria for lighting the feedback lights.

The technician sits at a control display panel adjacent to the plethysmograph and starts and stops the test from a keyboard. The subject's flow-volume curve is displayed for the technician while the test is in progress on a storage CRT scope so that he can monitor the subject's performance and detect bad tests, misunderstandings of instructions, and submaximal efforts. At the end of each trial the subject's percentage of normal VC and $FEV_{1.0}$ is also displayed on the scope.

The visual feedback display, shown in Figure 2, was described in the previous section. The computer calculated the subject's predicted normal values for peak flow and VC and lighted each light as its criterion flow rate or volume (70%, 80%, 90%, and 100%) was met in the test. The end of test light was ignited if the subject sustained his expiration for at least 0.5 seconds after the flow rate decreased to 0.45 liters per second.

For pneumatic calibration of the plethysmograph, a large syringe with a motor driven piston provided a known, constant volume of 4.3 liters and a flow rate of 3.9 liters/second. The equipment was calibrated to these known flow and volume rates at the start of each day's testing and rechecked at the end of each day. All electronic equipment was given approximately 30 minutes to warm up before calibration and testing began.

Procedure

Subjects were randomly assigned to one of four conditions: one experimental group performed five spirometry trials with the aid of the feedback display, a second experimental group received both verbal coaching from the technician and visual feedback from the display, one control group performed five spirometry trials without visual feedback but with the standard verbal encouragement from the technician during the test, and a second control group received instructions before the test started but no visual or verbal feedback, encouragement or knowledge of results were given during the course of the test.

The experimental groups were given instructions at the start of testing explaining that the lights are cued to their own performance and that they should try to inspire fully, exert maximal force in the blow-out maneuver, and maintain the expiratory effort as long as possible. They were instructed concerning what to do if they fail to light all lights of a given color. In addition, reminders were printed on the display under each set of lights: for flow lights--"To Light Blue Lights Blast Air Harder," for volume lights--"To Light Green Lights Take Deeper Breath," and for end-of-test light--"To Light Red Light Keep Pushing Longer." The lights group received no verbal encouragement or feedback from the technician after the start of testing. The combined feedback group received both coaching from the technician and feedback from the lights display.

The verbal feedback group was tested under standard spirometry procedures by an experienced technician. The procedure was standardized for all subjects: the technician talks and cheers loudly and continuously to the subject from the start to the end of each of the five trials. He exhorts the subject, "Take a deep breath, all you can hold, all you can hold," repeating this until the subject seems to have reached maximal inspiration, "now BLAST it out, push, push, keep pushing, keep pushing...." repeating this until the subject seems unable to expire any more air. At the end of each trial the technician tells the subject how well he has done and what he should do to make the next trial better.

The no feedback group was given the same standard instructions at the beginning of testing as the verbal feedback group, but was given no encouragement, feedback, or cheering from the technician during the course of the test or between trials. Subjects in all groups were tested for five trials of one forced expiratory maneuver each.

The test instructions given before testing were the same for all four groups. These were read aloud to the subject:

> This experiment is to test some of your breathing capacities. The test requires you to take in as big a breath as you can, blast it out as hard as you can, and keep exhaling until all of that breath is gone before you breathe in again.
>
> It is important that you breathe in as much as you can hold, perhaps straining a little at the start of the test. Than you should blast out that breath with as much force as you can muster. At the end of the breath be sure to exhale all you can before breathing in again.

The subjects in the lights feedback and combined feedback groups then received the following additional instructions concerning the display:

> To let you know how well you are doing this we have installed some feedback lights. The better you perform the test, the more lights you will light. The green lights will tell you how well you are doing at taking in a full breath. If all green lights are not on at the end of your test, you should concentrate on inspiring fully on the next test.
>
> The blue lights will tell you if you are blasting out the air hard enough when you start the forced exhaling. If you do not light all the blue lights, you should put more effort and force into your blast on the next trial.
>
> The red light will tell you if you are sustaining effort long enough at the end of the blow-out to make sure your lungs are

> completely empty. You should try to keep blowing out at least until the red light comes on to assure that there is no air left in your lungs.

For all groups, the following measures were calculated: (1) percentage of predicted normal forced vital capacity (PVC), (2) percentage of predicted forced expiratory volume in one second (PFEV1), (3) percentage of predicted peak flow (PPF), (4) flow rate at 50 percent of the vital capacity (V50), and (5) time of expiration, or time of vital capacity (VCT), which is the time in seconds from the start of expiration until the last of the breath is expired.

The best of the five tests taken on each subject was chosen based on the highest PVC. The values of the other dependent variables chosen for analysis were those occurring on the best PVC trial.

The data was analyzed statistically by first multivariate then univariate analyses of variance. An alpha level of .10 was chosen as an acceptable level of significance because guarding against Type II error becomes more important when testing a method which may aid in medical diagnosis. Falsely rejecting a method with potential health benefits on the basis of too conservative a test would be an improper balancing of Type I and Type II error risks. Especially in view of the relatively small number of subjects tested, probability levels between .05 and .10 are viewed as promising, indicating potential benefits of the method.

CHAPTER V

RESULTS

The design used for data analysis included one independent variable at four levels with five dependent variables. The levels of the independent variable, feedback, were: (1) no feedback, (2) verbal feedback, (3) lights display feedback and (4) combined verbal and lights feedback.

The data analysis included only 58 of the 60 subjects originally tested because two subjects' data was lost from magnetic computer tape storage due to a programming error. This left 15 subjects each in the no feedback and verbal feedback groups, and 14 subjects each in the lights and combined feedback groups. The means and standard deviations for each group on each of the five dependent variables are shown in Table 2.

A multivariate analysis of variance was performed to test the effects of feedback on a composite of the five dependent variables. The overall multivariate F was significant ($p < .001$). The only significant multivariate comparison was that no feedback was lower than all the other feedback conditions ($p < .01$), which were not significantly different from each other.

Univariate analyses of variance were then performed for each of the dependent variables. The alpha level was adjusted to .02 for the number of dependent variables. These showed significant overall F ratios for percent vital capacity ($p < .012$) and for the percent peak

Table 2. Mean and Standard Deviation for Each Feedback Group on Each Spirometric Measure

	FEEDBACK			
	None N=15	Verbal N=15	Lights N=14	Combined N=14
%VC	81.1 15.6	92.1 15.6	96.4* 10.7+	96.2* 11.9
%FEV1	88.8 17.3	99.3 19.3	102.4 9.2+	101.1 9.0+
%Peak Flow	78.4 19.0	105.8* 14.8	102.0* 12.0	105.1* 12.9
Time of VC seconds	2.52 .69	3.04 .96	3.55 1.60	3.65 1.42
$\dot{V}_{50}$ liters/sec.	3.96 1.08	4.50 1.08	4.74 .87	4.53 .90

*significantly different from no feedback
+significantly lower variance than for verbal feedback

flow ($p < .001$). No significant effects of feedback on $\dot{V}_{50}$ were found. Non-orthogonal multiple comparisons were made by Tukey's Honestly Significant Difference (HSD) procedure (Kirk, 1968) for each of the two dependent variables that had shown a significant overall effect.

For percent of predicted vital capacity (PVC) both the lights group and the combined feedback group were significantly better than the no feedback group, while verbal feedback was not significantly better than no feedback. The means for the four feedback groups were no feedback = 81.1, verbal feedback = 92.1, lights feedback = 06.4, and combined feedback = 96.2.

For percent of predicted forced expiratory volume in one second (PFEV1) the group means were: no feedback = 88.8, verbal = 99.3, lights = 102.4, and combined = 101.1. None of these differences were significant in the analysis of variance ($p < .06$) but large deviations from the homogeneous variances assumption called this result into question. An F max test showed that the homogeneity assumption was untenable ($F_{max} = 4.57$, $p < .05$). A nonparametric analysis was therefore performed for PFEV1. The Kruskal-Wallis test, an analogue to the analysis of variance, (Hollander and Wolfe, 1973) showed nonsignificant overall effects ($p < .10$).

Percent of predicted peak flow (PPF) was the only variable on which lights feedback did not produce a higher average than verbal feedback. The differences between the means for verbal (105.8), lights (102.0), and combined feedback (105.1) were not significantly different and all were significantly better than no feedback (78.4).

The time of the vital capacity (VCT), a measure of sustained effort, showed the following group means: no feedback = 2.52 seconds, verbal feedback = 3.04 seconds, lights feedback = 3.55 seconds, and combined feedback = 3.65 seconds. These differences were not significant in the analysis of variance, but again large group differences in variability (F_{max} = 5.39, $p < .05$) called the parametric test into question. A Kruskal-Wallis nonparametric test failed to show overall significance ($.05 < p < .10$).

Since the overall ANOVA for the fifth dependent variable, the flow rate at 50 percent of the vital capacity ($\dot{V}_{50}$), showed no significant effects, multiple comparisons were not made. The group means and standard deviations for $\dot{V}_{50}$ are shown in Table 2.

After statistical analysis of group means, large differences in group variances suggested that the use of the feedback display might significantly reduce variability as well as increasing mean performance. The variances of the lights feedback and the combined feedback groups were compared with the variance of the verbal feedback standard spirometry group. For PVC the variance of the verbal feedback group was 2.13 times greater than that of the lights feedback group ($p < .10$). For PFEV1 the variance of the verbal feedback group was 4.39 times that of the lights feedback group ($p < .01$) and 4.57 times that of the combined feedback group ($p < .01$). On PVC, PFEV1, PPF, and $\dot{V}_{50}$ the variability of the verbal feedback group was in all cases higher than both lights and combined feedback groups, but only those comparisons mentioned above were significant. The standard deviations for all groups on all variables are shown in Table 2.

CHAPTER VI

DISCUSSION

The hypotheses stated previously can be evaluated by means of these results. Hypothesis 1, that the two groups receiving lights display feedback would both perform better than either of the other groups on a composite of the respiratory parameters, was not confirmed. The only significant difference in the multivariate analysis was that all three groups receiving any kind of feedback were significantly better than the no feedback group. The verbal, lights and combined feedback groups did not differ significantly on the multivariate composite score.

Hypothesis 2, that the beneficial effects of the display should occur on the variables PVC, PFEV1, PPF, and VCT while $\dot{V}_{50}$ should not be affected, was partially confirmed. No significant effects were found for $\dot{V}_{50}$. This was expected since the flow rate at midexpiration is somewhat independent of effort (Dayman, 1967). On the other four dependent variables the lights display elicited higher values than no feedback, but differences were significant only for PVC and PPF. On PVC both of the groups receiving visual feedback were significantly better than no feedback while verbal feedback alone was not significantly different from no feedback. The hypothesis, however, cannot be unconditionally confirmed because verbal, lights, and combined feedback groups were not significantly different from each other on any of the dependent variables and no significant effects were found for PFEV1 and VCT.

Hypothesis 3, that the feedback display should reduce variability in PVC, PFEV1, and PPF, was confirmed for the first two measures but not for PPF. Variability among subjects was reduced to less than one-half that of the standard spirometry condition by the use of the display for percent vital capacity and to less than one-fourth percent FEV1. Variability was slightly, but not significantly reduced on percent peak flow and $\dot{V}_{50}$.

These results demonstrate clearly the usefulness of a visual feedback display in improving the quality of spirometric test results. The vital capacity is the most basic and most important pulmonary function measure for diagnosis of respiratory disease. That a visual feedback display cued to the subject's performance can produce higher values and lower variances on this variable suggests that its use could lead to more accurate diagnoses.

The dramatic decrease in variability when the feedback display is used also has implications for medical research as well as medical diagnosis. The incorporation of such a display into spirometric research would reduce error variance due to technician effects, subject motivation, effort, and comprehension of instructions, thus making the experiment more sensitive to treatment effects. That these variability effects occur with an increase in the mean values on all the respiratory variables suggests the benefits of using such a display on a routine basis in any computerized pulmonary function laboratory.

It should also be pointed out that several considerations suggest that the visual display effects may have been underestimated here. First of all, variability effects might have been even larger if the experiment had included more than one technician. The reduction of inter-technician effects by minimizing technician involvement and providing more constant test protocol through the use of a display should cause an even more dramatic decrease in variance than simply reducing intra-technician effects.

Secondly, several of the effects declared nonsignificant in the univariate analyses (PFEV1 and VCT) approached significance and might be found significant in an experiment which utilizes more than the relatively small (15 per group) number of subjects tested in this study.

Finally, all the subjects tested here were cooperative, healthy college students with above average intelligence. The feedback display might be even more effective with subjects in an occupational screening study or hospital patients who might be less cooperative or have more trouble understanding the instructions for the test. The interest-catching and instructional advantages of the display could become more important.

The value of such a feedback display is indicated for several reasons: (1) Decreased variability can lead to more accurate and reliable test results. (2) Inter-technician and intra-technician effects can be reduced. (3) Higher values on important respiratory measures can lead to more accurate diagnosis. (4) The role of the technician in the test can be minimized. In mass screening studies this can mean less fatigue and voice strain for the technician. (5) The testing procedure can be more standardized and controlled.

CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

The results showed that the use of the display yielded better mean performance on vital capacity and at least as good as standard spirometry on forced expiratory volume in one second, peak flow rate, the time of the vital capacity, and the flow rate at 50 percent of the vital capacity. A marked reduction in variability also resulted from use of the visual display. For the visual display group the within-cell variability on percentage of predicted normal vital capacity was less than one-half the variability of the standard verbal feedback group. On percent of predicted forced expiratory volume in one second both the display group and the display plus verbal feedback group showed less than one-fourth the variance of the verbal feedback group.

These results suggest that use of a display like this one as a standard part of a pulmonary function testing system would increase the quality and reduce the variability of the test results. This could lead to more accurate medical diagnoses based on the test and, in experimental studies, a greater sensitivity of the test to treatment effects.

This study also suggests several implications for further research. First, a larger sample might show significant effects of the feedback display on some of the respiratory variables that approached significance here (VCT and PFEV1).

Secondly, extending the design to include a technician factor should provide information on the benefits of the display in reducing inter-technician effects on test results, which were not considered here. Male and female technicians might be included to assess the usefulness of the display based on the sex of the technician.

Finally, since the sample used here was relatively young, healthy, and intelligent, information is needed concerning the effectiveness of the display on the two types of populations who are more likely to be the target populations for pulmonary function tests. The two populations are hospital or doctors' patients who are being diagnosed for respiratory disease and blue collar factory workers who receive occupational lung screening. These populations will be much more variable with respect to intelligence and health than was the sample used in the present study.

BIBLIOGRAPHY

Alluisi, E. A., and Morgan, B. B. Engineering psychology and human performance. Annual Review of Psychology, 1976, 27, 303-330.

Bilodeau, E. A., Bilodeau, E. McD., and Schumsky, D. A. Some effects of introducing and withdrawing knowledge of results early and late in practice. Journal of Experimental Psychology, 1959, 58, 142-144.

Block, J. D., Lagerson, J., Zohman, L. R., and Kelly, G. A feedback device for teaching diaphragmatic breathing. American Review of Respiratory Disease, 1959, 100, 577-578.

Brown, B. B. The biofeedback syllabus: A handbook for the psychophysiologic study of biofeedback. Springfield: Thomas, 1975.

Caplan, R. D., Cobb, S., French, J. R. P., Harrison, R. V., and Pinneau, S. R. Job demands and worker health. HEW publication No. 75-160, 1975.

Cohen, A., Smith, M., and Cohen, H. H. Safety program practices in high versus low accident rate companies--An interim report. HEW publication No. 75-185, 1975.

Dayman, H. G. Flow control during forced expiration--A preliminary report. American Review of Respiratory Disease, 1967, 95, 887.

Discher, D. P. LS/MFT: Lung screening by meaningful function testing--Part 1. American Journal of Public Health, 1970, 60, 2118-2132.

Discher, D. P., Massey, F. J., and Hallett, W. Y. Quality evaluation and control methods in computer-assisted screening. Archives of Environmental Health, 1969, 19, 323-333.

Discher, D. P., Massey, F. J., and Otoupalik, E. L. An evaluation of screening tests for chronic lung disease. Journal of Occupational Medicine, 1970, 12, 429-443.

Discher, D. P. and Palmer, A. H. Development of a new motivational spirometer--Rationale for hardware and software. Journal of Occupational Medicine, 1973, 14, 679-685.

Elwell, J. L. and Grindley, G. C. The effects of knowledge of results on learning and performance. British Journal of Psychology, 1938, 29, 39-54.

Henry, J. P., Junas, R., and Smith, K. U. Experimental cybernetic analysis of delayed feedback if breath-pressure control. American Journal of Physical Medicine, 1967, 46, 1317-1363.

Hollander, M. and Wolfe, D. A. Nonparametric statistical methods. New York: John Wiley, 1973.

Kirk, R. E. Experimental design procedures for the behavioral sciences. Belmont, Cal: Wadsworth, 1968.

Leiner, G. C., Abramowitz, S., Small, M. J., Stenby, V. B., and Lewis, W. A. Expiratory peak flow rate: Standard values for normal subjects. Use as a clinical test of ventilatory function. American Review of Respiratory Diseases, 1963, 88, 644-651.

Macpherson, S. J., Dees, V., and Grindley, G. C. The effect of knowledge of results on learning and performance: II: Some characteristics of very simple skills. Quarterly, Journal of Experimental Psychology, 1948, 1, 68-78.

Morris, J. F., Koski, A., and Johnson, L. C. Spirometric standards for healthy nonsmoking adults. American Review of Respiratory Disease, 1971, 103, 57-67.

Palmer, A., Ayers, W. R., Abraham, S., and Wilbur, M. Spirometry using motivational techniques. Journal of Chronic Disease, 1971, 24, 643-650.

Prather, K., Crisera, R. A., and Fidell, S. Behavioral analysis of workers and job hazards in the roofing industry. HEW Publication No. 75-176, 1975.

Repko, J. D., Morgan, B. B., and Nicholson, J. Behavioral effects of occupational exposure to lead. HEW Publication No. 75-164, 1975.

Ruppel, G. Manuel of pulmonary function testing. St. Louis: C. V. Mosby Co., 1975.

Sleight, R. B. and Cook, K. G. Problems in occupational safety and health. A critical review of select worker physical and psychological factors. HEW Publication No. 75-124, 1974.

Smith, K. U. and Henry, J. P. Cybernetic foundations for rehabilitation. American Journal of Physical Medicine, 1967, 46, 379-468.

Thorndike, E. L. The law of effect. American Journal of Psychology, 1927, 39, 212-222.

Trowbrigne, M. H. and Cason, H. An experimental study of Thorndike's theory of learning. Journal of General Psychology, 1932, 7, 245-258.

West, J. B. Respiratory physiology--the essentials. Baltimore: Williams and Wilkins Co., 1974.

Xintaras, C., Johnson, B. L. and de Groot, I. (Eds.) Behavioral toxicology. Early detection of occupational hazards. HEW Publication No. 74-126, 1974.