THE INFLUENCE OF KNOWLEDGE OF RESULTS AND GOAL SETTING ON PERFORMANCE WITH A MULTIPLE CHOICE ACADEMIC TEST

James Over Rust
New Kent, Virginia

B.A. Lehigh University, 1967

M.A. University of Richmond, 1969

A Dissertation Presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Education

School Psychology

University of Virginia

June, 1973

Approved by:

Brent Bridgeman
Assistant Professor
Chairman

Richard Abidin
Associate Professor

Russell Grieger
Assistant Professor

Harold Strang
Associate Professor

William White
Assistant Professor

ACKNOWLEDGEMENTS

This dissertation was made possible because of the generous and insightful help of my wife, Jane. She acted as both editor and typist, and provided encouragement and support when things got hurried.

My Committee Chairman, Dr. Brent Bridgeman, allowed me considerable freedom in developing my own ideas, but intervened offering sound advice when my thinking became confused. The fact that Dr. Bridgeman kept his office door open meant that he could be contacted when his assistance was most needed.

Dr. Harold Strang served as what amounted to Vice-Chairman of the Committee. The disseratation was an extension of research developed by Dr. Strang. His energetic criticism was a driving force behind this work.

I would also like to think Dr. Richard Abidin, Dr. Russell Grieger, and Dr. William White for serving on my Committee and providing conceptual and methodological recommendations.

Mr. Dan Keeling, Dr. Herbert Richards, and Mr. Pedro Saavedra offered statistical assistance which became essential when a new generation of computers arrived on campus.

The classroom teachers, Mrs. Pam Keeling and Mrs. Mary Anderson were enthusiastic and cooperative participants in the actual experiment. Their zeal enlivened their students' interest in the academic tests. I would also like to thank

the principal of the Orange Intermediate School, Mr. Ewart Robinsoh, for allowing me to use students from his school in the study.

Finally I would like to thank the Educational Testing Service for permitting me to reproduce the STEP Tests at a rate below purchase price of the test booklets.

ABSTRACT

Immediate knowledge of results (IKR) and goal setting (GS) instructions were administered to 304 seventh grade students in a academic testing situation. Neither of the independent variables was found to significantly influence test results. Supplementary analyses of the data called into serious doubt the actual functionality of the IKR and GS. The IKR failed to increase the students' ability to estimate their actual scores, and the GS instructions failed to alter their intentions. The possibility was raised that the testing situation was so anxiety provoking that the independent variables were masked. It was suggested that discrepancies in the IKR literature may have resulted from experimental procedures too weak to alter the mental sets of the subjects.

TABLE OF CONTENTS

LIST OF TABLES

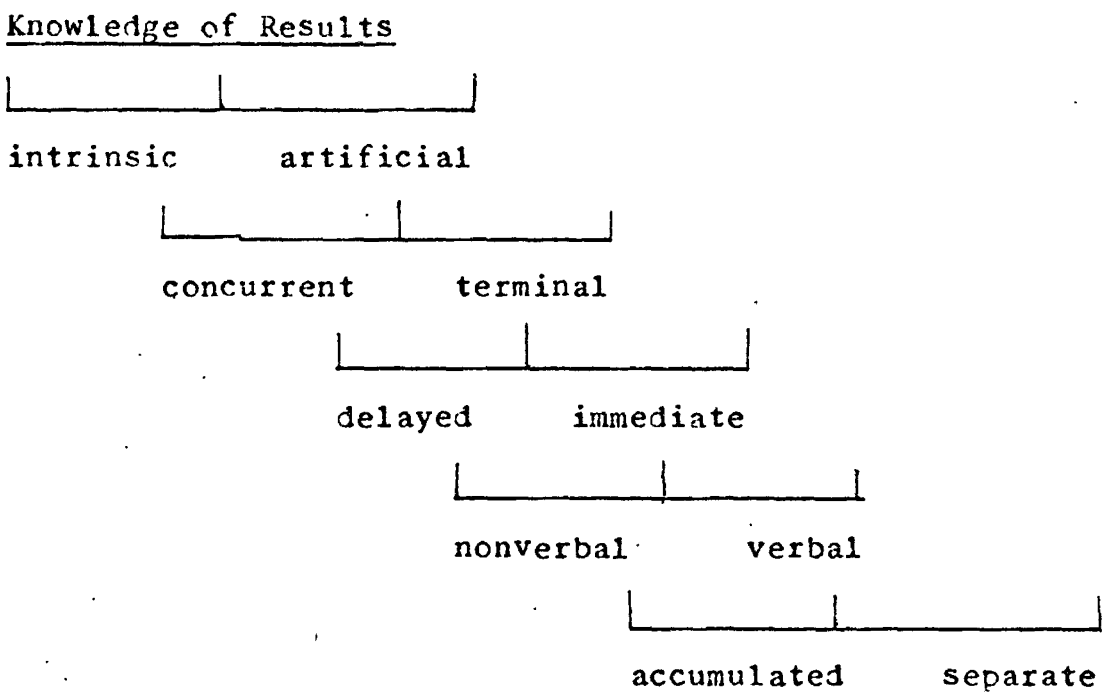
Chapter 1

GENERAL INTRODUCTION

Historically, knowledge of results (KR) and goal setting (GS) have been regarded as significant determiners of human learning and performance (Kirby, 1913). A review of the literature, however, shows a paucity of research that attempts to determine the separate and combined effects of these variables, particularly with children. KR has been found to provide a clearly beneficial effect on psychomotor tasks such as line drawing, ball tossing, and dynamometer squeezing (Thorndike, 1932). It has also been demonstrated to enhance performance on course related material for college students (Pressey, 1950) and on English pluralization rules for kindergarten pupils (Bryant and Anisfeld, 1969).

In studying the effects of knowledge of results there are a variety of definitional problems which must be addressed. "Knowledge of results" has been used to refer to information gained immediately upon responding (Angell, 1949), to total scores reported on previous work (Brown, 1932 and Berglund, 1969) and to scores reported at the end of every trial (Rucinski, 1968). Holding (1965, p. 22) has delimited knowledge of results by dividing it into eight dichotimous

classifications (only one branch is presented here):



Intrinsic KR is the feeling of correctness that comes to the subject who is performing a given task. A person will have some idea of how well he is doing because of his past history and his current interest. Artificial KR is under the control of the experimenter. Concurrent KR is exemplified by a meter which records ongoing behavior; terminal KR appears only after a response has been completed. KR may be delayed for seconds, hours, or days or it may follow directly after a response. Nonverbal KR might be the sight of a basketball rolling off the rim, while the verbal KR might offer corrective suggestions.

Separate KR is given individually after each response with no effort to accumulate it. Holding (1965) suggests that there is no reason why separate and accumualted KR can not be given. Payne and Hauty (1955) have added further clarification by suggesting "directive and incentive" as two additional types of knowledge of results. Directive KR provides the subject with cues which are useful in later performance, while incentive KR operates exclusively at the motivational level.

A variety of terms have been used to refer to KR, among them are "feedback," "information feedback," "knowledge of correct response," "immediate knowledge of results" and "knowledge of score." Due to the number of terms it is important to make clear exactly how one intends to use KR in a given situation. In this study the KR variable will be called "immediate knowledge of results" (IKR). Immediately upon responding, the subject will know whether he is correct or not. He will not, however, be given any information that will help him on other questions. IKR will be artificial because it will be supplied by the experimenter; it will be be nonverbal, and it will be separate for each response.

The existing evidence on the influence of goal setting on performance has been generally consistent. Goal setting has been shown to positively influence erograph pressure

estimations (Wright, 1906) and numerical cancellations (Anderson and Brandt, 1939). It has enhanced performance on programmed instruction with a normal population (Berman, 1967), and on spelling scores with a retarded sample (Warner and de Jung, 1971). Both Anderson and Brandt (1939), and Locke (1968a) found a linear relationship between goal setting and performance. Subjects who set high goals performed relatively better than subjects with low goals. This finding was replicated on tasks involving letter cancellation, object listing, complex computation, perceptual speed and addition (Locke, Cartledge, Knerr, and Bryan, 1969).

The influence of goal setting has been most powerful when the experimenter has been able to convince the subject to accept the assigned goal. Stedry (1960) warns that assigning goals after a subject has already expressed his choice is a weak procedure because subjects often do not change their intention. One way to avoid the problem is to allow the subjects to select their own goals and to be classified accordingly. A second solution is using a postexperimental questionnaire. It is acknowledged that subjects will set intrinsic goals whether or not they are instructed to do so. No effort will be made to control the intrinsic goals which the subjects may adopt.

Generally, knowledge of results and goal setting have been shown to increase the efficiency of the learning process. However, the roles of knowledge of results and goal setting in academic testing have not been clearly delineated. Anecdotally Terman (1937) suggests that psychological examiners ought to pause between subtests to offer encouragement to the testee regardless of the responses. Likewise, Wechsler (1949) firmly states that disapproving comments from the tester are to be avoided, and that praise should not be excessively used because "children vary in their reaction to commendation from an adult" (Wechsler, 1949, p. 19). In an experimental setting, Sweet and Ringness (1971) used the Wechsler Intelligence Scale for Children and found that while lower class while children's scores were increased by IKR, neither middle class white nor lower class black children showed any significant differences. Zontine, Richards, and Strang (1972) tested lower socioeconomic status children with the Peabody Picture Vocabulary Test. Three testing conditions were used: no KR, IKR, and IKR + reinforcement. Neither IKR nor IKR + reinforcement was found to significantly influence performance. Dahle and Daly (1972) found that retarded children's scores on the Wepman Auditory Discrimination Test were not significantly improved by providing feedback.

Providing feedback on group academic tests has also met with mixed results. Heald (1970) found that corrective feedback improved test scores in a classroom. On the other hand, Bierbaum (1965) found that IKR was related to inferior performance, while Gray (1968) and Montor (1970) found no significant KR vs. no KR differences on such test performance. Interpretation of the results of these studies is complicated by dissimiliar experimental designs.

Locke, Cartledge, and Koeppel (1968) showed that goal setting and KR have been confounded in many studies. For example, Kirby (1913) told children how well they were doing (KR) and encouraged them to improve on their prior score (GS). It was not surprising that the children's performance improved, but it is difficult to say exactly why. According to Locke et al (1968) goal setting is the critical variable. Sauer (1970), on the other hand, found data which contradicted Locke.

Socioeconomic status, sex, and anxiety level have been shown to influence the way people react to test situations. Sweet and Ringness (1971) found an interaction between SES and knowledge of results. Hakanson, Willers, and Koropsak (1968) discovered consistent sex differences in subjects' reactions to stressful decision making situations. Campeau (1968) found an anxiety level by KR interaction for girls but not for boys.

SES, sex, and anxiety level are all potentially relevant independent variables for the study of test performance.

Knowledge of results and goal setting are two areas which have been of great interest to psychologists and educators. This interest has spanned half a century and has provided some solid findings in the area of instruction. The aim of this research is to study the influence of KR and GS in an academic testing situation. It is expected that each of the main effects will significantly influence performance. It is hoped that this study will help to clarify the influence of KR and GS in a practical testing situation. If test performance is enhanced or hindered by immediate knowledge of results and goal setting, then psychologists and educators would be well served to know how.

Review of the Literature

Knowledge of results (KR) and goal setting (GS) have been studied separately and jointly. The studies reviewed here have been selected to reflect work from four areas: KR alone, GS alone, KR-GS together, and KR-GS and socioeconomic status, sex, and anxiety level combined.

KR Alone: Knowledge of Results Benefiting Performance. The influence of KR on performance has been shown to depend on whether the KR serves a directive or incnetive function. Ammons (1956) stated that "knowledge of performance affects rate of learning and level reached by learning" (p. 283). He was referring to directive knowledge of performance as typified by a series of experiments carried out by Thorndike (1932). In one the subjects tried to squeeze a dynamometer in a sequence of specific intensities. After twenty-four no KR attempts at each intensity, the experimenter aided the subjects by saying "right" or "wrong", but the subjects were not told the direction of their errors. Thorndike reported that his experimental subjects manifested rapid gains. Percentage correct rose from a mean of 18 without KR to a mean of 52 with KR. When KR was removed, the percentage correct fell to 36.5. Thorndike (1932) also reported a

blind-ball-toss experiment. He found that the "right" - "wrong" KR group displayed an average superiority of 218% over the no KR group.

The effect of knowledge of results on programmed instruction has been well researched. Here again KR is clearly corrective, and is providing instruction. Only a few studies will be presented here. Anderson, Kulhavy, and Andre (1971) published two experiments which demonstrated that subjects who received knowledge of correct response (KCR) after each frame of a programmed lesson performed better on a criterion test than subjects who were not given KR. Eight feedback conditions were compared on a computer based program on the diagnosis of myocardial infraction. In their second experiment Anderson et al (1971) added a "peek" group. In this condition KCR was presented with the frame but the subject was warned not to sneak a look at it. This condition was similiar to arrangements of standard programmed instruction and allows the subjects to follow the law of least effort (Anderson, 1970), by peeking at the answer before he thinks about the question. The "peek" group did poorer than any other group on the criterion test. The no KR group along with the "peek" group did significantly poorer than the KCR group.

In other experiments which provide peek information, Krumboltz and Weisman (1962) varied the amount of KR on a

programmed text and found a significant negative linear trend between percentage of confirming statements and errors on the program. They failed to find any differences on a criterion test. Gray (1968) also found no difference on a criterion test which followed four KR vs. no KR pretests. He notes that "a significant number (of subjects) expressed frustration at knowing that they had missed items" (Gray, 1968, p. 9). These studies show that when dealing with collective KR the critical factor is when the KR is made available. If the KR is presented after a response, then it has been shown to lead to improved performance. However, KR might actually harm later performance if it is made available prior to a response.

Pressey (1950) has developed a form of programmed instruction which is of particular interest here. "Each student has a set of multiple choice questions, sometimes supported by other material, a lecture, film, or text, and a device which automatically scores his answers, right or wrong. As an example, a four alternative choice question is answered by pushing a stylus into one of four holes in a board. Wrong answer holes are partly blocked and the stylus can only be pushed home in the right answer hole" (Annett, 1969, p. 97). The student is instructed to keep trying until he selects the correct choice. Pressey, and VanValkenburgh, Nooger, and Neville, Inc. have

developed more manageable feedback devices. These response cards provide either right-wrong KR or refer the student to appropriate pages in the text. They call the system adjunct auto-instruction. This system offers the advantage of controlling for cheating by eliminating peeking.

Pressey and Kinzer (1964) tested their auto-adjunct tutoring device with Holland and Skinner's (1961) programmed text, The Analysis of Behavior. Five groups of subjects were used. Group 1 was a control group which was given the criterion test without having read the text. The other groups either followed the Holland and Skinner procedure, read a succinct statement developed by the experimenters, read Holland and Skinner with all the answers provided, or read the succinct statement and answered fifteen auto-instruction questions. The study found that the Holland and Skinner suggested procedure took longer and taught less than the auto-instructional technique as measured on an essay type criterion test. Dowell, Gurney, and Norris (1956) demonstrate how the auto-instructional devices have been used by the Air Force to teach electrical trouble shooting. Finally, the influence of IKR was tested by Montor (1970) using Navy Midshipmen. IKR was provided to one half of the subjects via Pressey type response cards. The correct answers were read to all the subjects upon completion of each of twelve quizzes. The performance curves for the two groups

were remarkably similiar. There was no difference between average scores on the criterion test given three weeks later.

In a widely quoted study Kaess and Zeaman (1960) studied the use of a Pressey punchboard by varying the number of alternative responses on a multiple choice exercise. One group was forced to make the correct choice by being offered only one selection, a second group was offered two alternatives, and so on up to five. There was also a "5 alternatives" group which was not given knowledge of correct response. Learning was measured by parallel forms of the original five choice test. The groups' learning curves differed greatly. The group which was offered only one selection learned the fastest. Each of the other groups followed according to the number of alternatives originally offered. The five alternative, no KR group demonstrated the slowest learning. Kaess and Zeaman interpret these findings to mean that when subjects respond incorrectly, they have in a sense learned the wrong answer. This is so even when KR is eventually provided. It is strongest, however, when no KR is provided. The result of this false learning is that it must be undone. Relearning takes time and is reflected in the learning curves.

Karraker (1967) continued this line of research when he studied a group of college freshman. Some of them were exposed to a multiple choice pretest with KR: a second group was given

the same test but without KR; and a third group did not take the pretest. On a completion type criterion test, the "pretest but no KR" subjects recalled significantly more plausible-but-incorrect alternatives than either other group. The author also discussed the fact that the KR group did not recall the faulty alternatives. This was interpreted to mean that as long as KR is provided, multiple choice items can be made part of programmed instruction.

School room experiments have attempted to demonstrate the importance of KR on retention. In studies of this type KR serves a directive function since the KR provides instruction which is measured later. Frase (1967) assigned prose material and retention questions to college students. Upon finishing the assignment, the subjects were retested on the retention items. Subjects who were provided with KR during the task performed significantly better than subjects who were not given KR. Angell (1949) gave three mid-term quizzes to college freshman chemistry students. The experimental group was given immediate knowledge of results (IKR) by punchboards while the control group was tested with standard IBM answer sheets. The author did not compare the quiz grades of the two groups. On the final exam the IKR group out performed the no IKR group.

In a similiar study Paige (1966) investigated eighth grade students who spent four weeks learning a new numeration system.

During this period four quizzes and one test were given. Half of the students received KR when they handed in their work. The KR was provided by a special carbon treated answer sheet which showed both the student's response and the correct answer. Unfortunately quiz scores were not reported. The criterion test was given three weeks later, and on it the KR group significantly out performed the no KR group.

Experiments have been carried out which compared two types of corrective KR. These studies determined whether giving right-wrong KR was as powerful as giving KR plus corrective instruction. Travers, Van Wagenen, Haygood, and McCormick (1964) taught fourth, fifth, and sixth graders German vocabulary words. Half of the children were directly involved in responding to the flash cards. The experimenter taught by: (1) giving right-wrong KR, (2) giving only wrong KR, (3) giving right KR and correcting wrong responses, and (4) only correcting wrong responses. Groups 1, 3 and 4 all did significantly better than group 2, and group 1 outscored group 4. Gilman (1969) studied a computer assisted program dealing with basic science concepts. He found that telling a student what the correct answer should be or why a response was wrong, was more effective than right-wrong KR.

Ammons (1956) implied that non corrective feedback would influence performance when he stated that "the most common effect of knowledge of performance is to increase motivation" (p. 285).

If KR were non corrective, increases in performance associated with KR would be incentive related. Such a motivational effect was found by Hurlock (1925). She gave a series of math tests to grade school children. The control group was tested and retested four times alone. The experimental groups were placed together and systematically praised, reproved, or ignored. These comments were not contingent upon anything the children did. The results indicated that the praised children gained the most, then came the reproved, the ignored, and control groups. These differences were exaggerated when the children were divided into superior, average, and inferior ability groups. The effect of the verbal incentives was greatest for the inferior students. Manzer (1935) also found a clearly motivational effect. He tested 128 subjects for fifty trials on a hand dynamometer, and found that providing knowledge of output resulted in a sudden surge in muscular work.

Flook and Saggar (1968) and Means and Means (1971) found that giving subjects knowledge of scores on aptitude tests influenced subsequent academic performance. Since aptitude tests have no corrective influence on future academic performance, their impact was motivational. Flook and Saggar (1968) gave intelligence, aptitude, and personality tests to first year engineering students. Two matched groups were formed with the

experimental group being given scores from the tests in a truthful but encouraging manner. The control subjects were not given their scores, but were granted a personal interview to control for any motivation due to the attention. At the end of the year, the group given KR had out performed the no KR group. When the groups were subdivided by achievement into thirds, it was found that the influence of KR was greatest for the lowest third of the class.

Means and Means (1971) gave a psychology class a phoney "aptitude test for adolescent psychology." They then divided the students into high grade point average (GPA) and low GPA groups. During an individual conference held later in the semester, the students were given randomly assigned KR. They were told either that they had high aptitude for the course, low aptitude for the course, or the topic was not brought up and nothing about aptitude was discussed. A significant interaction between GPA and "aptitude" scores was found with high GPA students who had been assigned low scores doing better than high GPA students who had been given high scores. Low GPA students assigned low scores tended to do worse than low GPA students assigned high scores. While these results indicated that incentive KR influenced later performance, the relationship between knowledge of aptitude and performance seemed to be confounded by ability level.

The influence of incentive KR on ongoing test performance

was investigated by Morgan and Morgan (1935) when they studied a self-scoring testing technique which provided IKR. Two groups of introductory psychology students were given a true-false test four times with or without KR. Since these tests were given over two semesters, they presumably did not count toward a grade. In this study the influence of KR was measured twoce: first as an incentive influencing immediate performance, then as a corrective agen influencing subsequent performance. The second measure yielded clearer results. Although the groups did not differ significantly on the initial test, they did differ significantly on each subsequent retest. KR was thus found to be instructive as related to subsequent performance but not as relating to ongoing performance. In a similiar study Heald (1970) divided education students into three KR groups. Students in the first group (KR-R) had correct responses confirmed, and were referred to an appropriate page in the text to remediate incorrect responses. Students in the second group (KR) learned immediately whether their answers were right or wrong but no corrective passages were suggested. Students in the control group (C) received a standard multiple choice answer sheet but no corrective reading suggestions were offered. Heald found on a midterm that counted toward the course grade, that the KR-R group significantly outscored the C group. The KR group appeared to outscore the C group though not significantly.

On a retest both KR-R and KR groups performed significantly better than the no KR control group. A problem with this study was that in their remedial text assignments the KR-R group may well have learned something which aided them in other questions. A second problem was that careful examination of Heald's questions revealed possible crossover between items. Thus knowing the answer to one question may have directed a student toward correctly answering another. If there was information crossover, the KR becomes informational rather than motivational.

The studies reviewed demonstrate that directive KR has proven helpful in (1) increasing psychomotor accuracy, (2) increasing muscular motivation, (3) improving retention in programmed instruction, and (4) improving scores on related criterion measures. There was also a suggestion that incentive KR alone might assist subjects in ongoing test performance, but that such KR must be checked to be sure that it is non corrective. Knowledge of correct response which was made available before the subject responded was shown to detract from performance on a later test. Such knowledge offered the students an opportunity to complete the requirements of the lesson without putting forth much effort.

KR Alone: Immediate Knowledge of Results Relating to Inferior Performance. Evidence of situations where KR hurts performance is not extensive. Some researchers have found, however, that immediate KR is harmful compared to delayed KR. Brackbill, Wagner, and Wilson (1964) investigated differences between performance scores influenced by immediate KR (IKR) and ten seconds delayed KR. The dependent variable studied was acquisition of French vocabulary words by third grade boys. The findings were that subjects who received IKR retained fewer French words than subjects who had ten second dealyed KR. The retention measures were made one and seven days after the two learning sessions. Sassenrath and Yonge (1968) compared the influence of giving KR as soon as the task, sixty-four psychology questions, was completed with delaying such information for one day. Although no significant group differences appeared in performance tapped during or immediately after the program, the delay KR group showed significant improvement over the immediate KR group on a test given five days later.

Gutherie (1971) studied IKR vs. delayed KR and found similiar results. His subjects were college students who read academic type passages and took a completion test (not for credit). While one group of subjects got no KR, another group was given IKR, and a third group received twenty second

delayed KR. Retention tests were given either as soon as the subjects finished reading the passages or eight and a half minutes later. The delayed KR group retained significantly more than the IKR subjects. Both of the KR groups scored significantly higher than the no KR group. Gutherie was also interested in assesing the influence of KR on motivation. His dependent variable was the amount of time subjects would continue reading passages similiar to those they had just completed. He found that IKR subjects read significantly longer than delayed KR subjects. However, he found no idfference in his motivational index between the no KR and delayed KR subjects. Gutherie notes somewhat circularly, that subjects receiving IKR felt positive affect when given IKR, and that affect became associated with the passages. Thus when similiar passages were presented to them, they read with interest. Negative feelings caused by delayed KR caused those subjects to more quickly put aside thier continuation passage. The no KR group was reported as being somewhere in between.

An instructional program dealing with military justice was used by Sullivan, Baker, and Schutz (1967) to study the relationship between IKR and performance. During the course of instruction, eleven mastery quizzes provided either IKR of no KR. The no KR condition resulted in significantly higher

mastery scores and slightly (but not significantly) higher criterion test scores. These authors explained the differences by noting that subjects in the IKR condition used the mastery test for instructive purposes. It is apparent that these subjects found it easier to learn by answering questions and receiving feedback than by reading and studying the text.

Clark and Greenbery (1971) argue that KR theoretically hurts performance when the dependent variable being studied is stimulus recongition memory. A positive response to one of the distracting stimuli which is familiar but not part of the stimulus list just memorized results in corrective KR. Such a failure causes the subject to adopt a careful attitude and to set higher reporting limits. To make matters worse, when the KR subject fails to report a stimulus, he finds out that he is wrong. This type of a failure causes him to guess more freely and to lower his reporting limits on subsequent trials. The net result is a wide variation in cirterion reporting limits which causes a performance decriment. These experimenters studied thrity-two college students memorizing eighteen consonant-vowel-consonant (CVC) trigrams. Subjects were given three recognition test on which they had to distinguish previously learned trigrams and distracting "new" ones. The subjects were divided into KR vs. no KR, stressful directions vs. nonstressful directions. The results indicated that KR significantly

hindered the performance. When results of the stress indicator (an adjective checklist) were included the findings became even clearer. "Stressed subjects given knowledge of results were much more anxious than were the remaining groups. It is particularly instructive to contrast the performance measures of the highly anxious stressed / knowledge of results group with those of the less anxious unstressed / no knowledge of results group. According to the Studentized range statistic (Winer, 1962), both d' (the memory score) and L_x (the reporting limit) are significantly lower for the stressed / knowledge of results group..." (Clark and Greenberg, p. 44).

Recently Strang (1972) explored two independent variables: test vs. exercise and IKR vs. no IKR on influencing recognition on a light monitoring task. An a posteriori analysis revealed that IKR complete with "test" task definition harmed the performance of potentially high scorers.

Bierbaum (1965) studied the effects of immediate knowledge of performance on multiple choice test accuracy. College students served as subjects and scores on an examination that counted toward the final grade in a general psychology course constituted the dependent variable. To control variability, each subject received half of the test under the IKR condition and half the test under the no KR condition. Group 1 had KR

on the first part of the exam, while group 2 had KR on the second part. The experimental findings were that IKR scores were significantly lower than no KR scores. Anecdotally subjects reported that they felt under pressure in the IKR condition.

Results similiar to Bierbaum's were found by Strang and Rust (1973). In a 2 x 2 design, college students were divided into KR vs. no KR and test vs. exercise groups. All subjects took part A of the test under no IKR conditions. These scores were used as a covariate in the analysis of part B scores. It was found that significantly more errors were made by subjects in the IKR conditions. Students in the test conditions took significantly longer to take part B. On a nervousness questionnaire, both presenting the task as a test and giving IKR were signficantly related to increased nervousness. The authors concluded that using immediate feedback in assessment must be questioned.

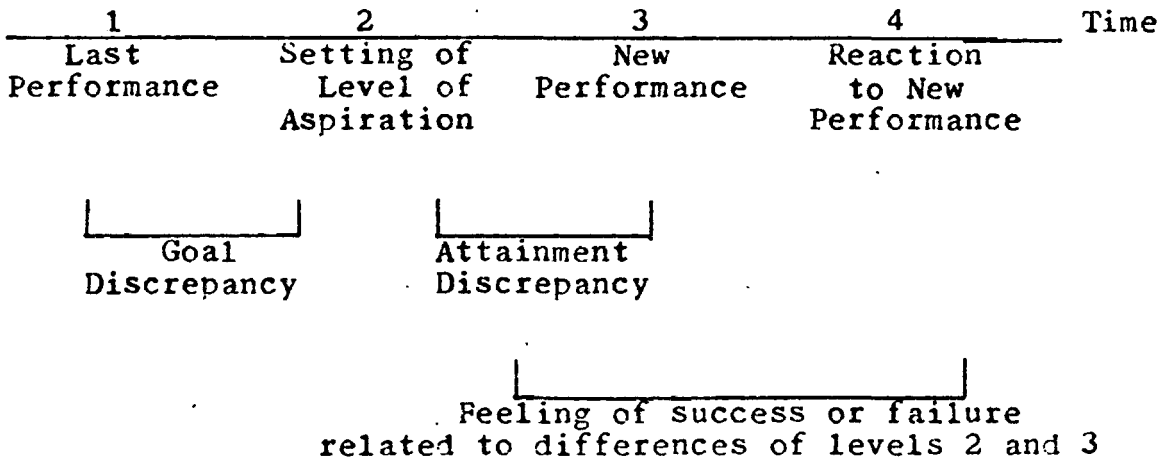
It would appear then that in some situations KR accelerates performance, while in others it depresses performance. To approach this discrepancy IKR must be viewed less descriptively and more functionally. If KR is corrective, it will improve performance so powerfully that it will mask the potentially important impact of incentive KR. If incentive KR is examined alone, it may be found to actually deter performance.

A second consideration is whether a study measures the influence of KR upon ongoing or subsequent performance. Measures which reflect ongoing performance are more sensative, and are likely to show motivational influences.

The availability of the KR relative to the response is also an issure. Studies have shown that in some instructional situations KR that is always present is detrimental. To summarize, while the facilitating influence of KR on ongoing and subsequent performance has been well documented in the literature, the influence of motivational KR upon performance is not fully understood.

Goal Setting: Theoretical Orientation. Ryan (1970) asked whether "intention (along with similiar terms such as task, purpose, desire, and goal) is a legitimate and useful explanatory concept to be used in psychology" ...and "if intention is used as an explanatory term, what is its relation to other terms and concepts used in explaining behavior (p. vi.) ?" Ryan answered his first question by stating his belief that: (1) experiences such as desires and goals can not be broken down, and (2) even if they could be divided they are still worth studying because they do influence behavior.

An attempt to partially answer the second question was made by Lewin, Dembo, Festinger, and Sears (1944) when they investigated level of aspiration (note Figure 2).



Typical Time Sequence in the Level of Aspiration Situation

Figure 2
From Lewin et al, 1944.

These authors assume that setting a goal is an important event which influences subsequent performance. They presuppose that a series of goals go together to make up level of aspiration. There are two levels of goals: the ideal goal and the action goal. The ideal goal is an individual's notion of perfection and acts as a general context for level of aspiration (LOA). In some ways the ideal goal resembles Ryan's (1970) "determining

tendency." Ryan used the term to refer to "the carry-over of intention to later behavior (p. 95)." The term was necessary to describe the situation when "at the time of appearance of the stimulus, there was no awareness of the task (intention or goal). The response simply came in a sort of automatic way, and the subject was only aware that the response was correct. In other words, the experience of intention was not immediately responsible for the reaction (p. 927)." Ideal goal and determining tendency refer to theoretical constructs which have developed in the LOA literature.

The action goal is more realistic and more data based. It is measured at point 2 on the diagram. Using goal discrepancy as dependent variables Lewin et al (1944) established a variety of hypotheses about LOA. The nature of the reference groups that the subject is compared with influences his LOA. The amount of recent success, the cultural norms, socioeconomic background, educational level, type and size of incentive, personality traits and individual differences were all cited as influencing goal setting. Cofer and Appley (1964) point out that situational factors such as subjective expectancy of success or failure play a critical role in level of aspiration.

Bryan and Locke (1967a) take Cofer and Appley's statement one step further when they say: "effort (or work rate) is

adjusted to the perceived difficulty of the task undertaken (p. 260)." The situation determines the expectancy which leads to a goal which influences performance. But there has been some controversy over the relationship between motivation and performance. Atkinson (1957) and McClelland, Atkinson, Clark, and Lowell (1953) have hypothesized an inverted "U" between strength of motivation and probability of success. They found that too little or too much motivation was harmful to performance. Eason (1963) studied rotary tracking and found an inverted U-shaped relationship between performance and the size of the target which in Ryan's (1970) terms was related to goal.

McClelland (1961) noted that high need for achievement (n Achievement) is positively related to performance (i.e., a linear relationship) on tasks which require thinking, but not on duller mechanical tasks. High n Achievement subjects improved their performance on mathmetics problems as they were led to believe that their chances of winning were becoming smaller. When faced with the same decreasing odds, low n Achievement subjects did more poorly at first and then improved their scores. The functional relationship between task difficulty and performance depends upon the subjects n Achievement rating.

Bryan and Locke (1967a) concluded that goal setting mediates performance. Two groups of subjects were given different goal setting instructions on an addition task, and were found to perform at significantly different levels with high goal subjects performing better than no goal subjects. Lee (1968) reviewed the literature in the level of aspiration area. He concluded that, among other things, LOA influences performance. He also concluded that there was a linear relationship between level of aspiration and ego involvement, and that there is a linear relationship between goal expectancy level and persistence.

Goal Setting and Knowledge of Results Studied Together: Experimental Settings. Frank (1935) mentioned KR in his definition of level of aspiration: "The level of future performance in a familiar task which an individual, knowing his level of past performance in that task, explicitly takes to reach (p. 119)." KR and LOA have been intertwined in the literature. Hertzman and Festinger (1940) studied the effect of KR and awareness of a reference group's score on subsequent goal estimates. A synonym and an information test were administered to college students. The subjects were given knowledge of their scores after each trial. They were also given knowledge of a reference group's average score and goal. A significant

influence on LOA was found. Individuals changed their aspirations to conform with those of their reference group. Anderson and Brnadt (1939) studied the relationship between goal setting and knowledge of results for fifth graders on six cancellation tasks. An experimental group was given KR and asked to set a goal. A control group simply repeated the task, looking for different numbers on each trial. The experimental group significantly outscored the control group, and a correlation of .57 was found between achievement and goal scores. Dey and Kaur (1965) found similiar results in their investigation of Indian female graduate students. It is interesting to note, however, that these experimenters found a drop in performance as goals reached the highest level.

Motivational techniques were ranked in terms of effectiveness by Bayton (1948). It was found that having each subject set his own goal, and then informing the experimenter was the most effective technique. Furthermore having subjects set their own goal without informing the experimenter, or having the experimenter give KR but no goal, were less effective. The least effective technique was no goal and no KR.

Fryer (1964) studied acquisition of skill in receiving International Morse Code and reached opposite conclusions. Independent variables included KR alone vs. KR with goal

setting, public expression vs. private expression of goal, low difficulty vs. high difficulty characters, and expect goals vs. hope goals. He found that when the characters were combined across difficulty levels, or just difficult characters were considered, that the goal setting groups significantly out performed the KR alone group. The other independent variables were generally insignificant. Fryer concluded that his study clearly demonstrated a definite superiority for goal setting.

Locke (1966b) reanalyzed Fryer's findings and came up with a somewhat modified interpretation. Locke stated that the level of the goal, not just the fact that a goal was set, was the critical factor. Subjects who set low goals did not perform differently than the subjects who set no goal. On the other hand, subjects who set high goals ended up learning more. Locke logically argues that the goal setting preceded performance changes. He determined that in thirty-four of thirty-nine relevant cases mean improvement scores were in the expected direction. He compared instances where LOA scores exceeded immediate past performance with instances where LOA was below prior performance. Locke concluded the paper by suggesting that postexperimental questionnaires be used to determine goal level of all subjects including those not asked to set a LOA.

More important to the present study are investigations

which used goal setting as an independent variable. Kausler (1959) hypothesized that performance on a simple task would improve as a result of having subjects express goals. An analysis of covariance found a significant goal setting main effect. A second experimental group significantly out performed the control group but was not different from the first experimental group.

Stedry (1960) studied performance on Luchin's water-jar problems. He varied the level of efficiency required for reinforcement (the number of problems which had to be solved in a set amount of time in order for the subject to get paid up to $9.00 per test). He called these levels "budgets". In his experimental groups of low, medium, and high, he varied the budget requirements for each trial and told the subjects what they had to do to be paid. The control group was under "implicit budget control." They were not informed except by after the fact payment. For some subjects goal setting (aspirational level) was measured before budget information was given. For other subjects goal setting was measured after they had received budget information. One third of the subjects stated no aspirational level. Significance was found for both the budget main effect and the goal setting main effect. Implicit budgets rated as more effective than medium, high, and low budgets in that

order. Both goal setting groups out performed the no-goal setting group. The study demonstrates (as did Fryer, 1964) the powerful influence of goal setting when all subjects were receiving KR.

In a series of experiments Locke, Cartledge, Knerr, and Bryan (1969) stressed the importance of rigorous goals. Beginning with the theories of Ryan which subsequently have been published (Ryan, 1970), the relationship between intention (goal setting) and performance was studied. In three experiments Locke (1966a) concluded that the level of performance depends directly and linearly upon the level of intention. In experiment 1, subjects were given KR and specific goals. They worked on a task of naming appropriate nouns which could be modified by a given adjective. In experiments 2 and 3, subjects were either given goals, or selected their own. Here the task involved naming uses for objects. A definite linear trend was found in each instance. The relationship between GS and performance was strong and positive.

A complex coordination task was the dependent variable for a further study (Locke and Bryan, 1966b). Subjects were either given specific goals which they were to reach in order to be successful or told to do their best. A postexperimental questionnaire was administered to check on whether the instruc-

tions were followed. Results indicated that goals helped performance and that instructions were generally followed. These same investigators further strengthened their position by studying new tasks (number cancellation and addition) over extended time periods.

Bryan and Locke (1967b) studied the influences of goal setting on lowly motivated subjects' subsequent performance. Motivation was measured by (1) the number of simple addition problems completed during a pretest, and (2) by ratings on a interest-boredom scale. The low motivation students were placed in the goal group while high motivation students were told to "do their best." The goal group was given specific goals for each page of problems. The treatment was found to significantly influence indicies of effort, performance and interest.

Locke, Bryan, and Kendall (1968) introduced cash incentives as an additional independent variable. Their tasks included giving uses for objects, toy construction, and word unscrambling. The authors concluded that goals related to performance and that when the goal influence had been partialed out, the influence of cash incentive was not significant. These studies have stressed the importance of goal setting as influencing motivation and performance.

Goal Setting Studied in the Classroom While KR is Held Constant. A few studies have investigated the influence of goal setting on course related material. Berman (1967) required that subjects achieve increasingly more difficult levels of performance in order to receive points and found that he could shape the performance of college students on programmed instruction. Point contingencies controlled either rate or accuracy depending upon the situation that had been established.

Freshman engineering students with high academic competence were studied by Uhlinger and Stephens(1960). An intentional measure of goal setting was found to predict performance. Battle (1966) studied junior high students and found that math and English grades were significantly related to goals. Unfortunately, since these studies were carried on late in the semester, the results could have been confounded by intelligent awareness of marks which contributed heavily to the actual final grade.

Locke and Bryan (1968a) correlated grade goals, which were made early in the semester, with actual performance. Correlations were taken for each subject's easiest course, hardest course, and for a common history course. Correlations were found to be highly significant ($\bar{r}$ = .39). This finding was taken to indicate that conscious goals influence performance in real life situations.

Educable retardates from a public school special education class and from a residential setting were given ten spelling tests to determine whether goal setting was an effective motivator (Warner and de Jung, 1971). The facilitating influence of goal setting was affirmed for both groups.

If it is assumed that goal setting influences performance, the question becomes, how to influence goal setting. Gaa (1970) used goal setting conferences with first to fourth graders. Three experimental groups were compared. Group 1 was taught hard goals, group 2 spent the conference time learning about studying, and group 3 spent the time in reading instructions. The dependent variables were reading scores and attitudes toward reading. The conferences were found to influence attitude, but reading scores significantly changed for only the first and second graders. The author suggests that students in group 1 became more realistic in their aspirations as a result of the treatment.

The Importance of Goal Setting vs. Knowledge of Results. Some controversy has developed over the relative importance of GS and KR when they are combined. Locke, Cartledge, Knerr, and Bryan (1969) conclude that "knowledge of results does not affect task performance independently of its effect on the individual's performance goals (p. i.)." Sauer (1970), however,

found that KR influenced performance while goals did not. He also found that KR did not influence intentions; thus KR operated independently of goal setting. To complicate matters even further Cummings, Schwab, and Rosen (1971) turned the question around and found that both KR and past performance influence goal setting. Schematically the situation is this:

Locke: KR + many other things -> GS -> Performance

Sauer: KR -> Performance -> GS
↘ GS

Cummings: Past Performance + KR -> GS

In a second series of studies, Locke and his associates concluded that goal setting mediates the influence of KR and independently affects performance. Their position was based on a number of experiments using a college population. Locke and Bryan (1966a) asked subjects to perform complex computations under three conditions. All subjects were given knowledge of trial by trial correctness, while score keeping was varied. The subjects either had knowledge of score (KS) plus goal setting instructions, KS alone, or no KS and no GS. All subjects were asked to specify their intentions on a postexperimental questionnaire. These responses demonstrated that the students were generally not assuming the goals which they had been assigned. The experimenters, therefore, redivided the subjects

according to their own goals and found significant performance differences. KS was not found to have any influence on performance except in how it influenced goal setting.

In a second experiment Locke and Bryan (1968b) employed the same complex math task, but administered intentional questionnaires after half of the trials as well as at the end. The results indicated that performance improvement and goals were positively correlated ($r = .57$) but when GS influences were mathematically removed, the relationship of KS to performance was not significant. Cummings, Schwab, and Rosen (1971) argued that the influence of performance on the first half of the trials was the critical factor. Thus prior performance was viewed as influencing subsequent performance with GS simply reflecting a subject's past performance.

Locke (1967) used KR and GS as independent variables to test for their influence directly. Simple addition was again the task studied. Subjects were divided into four matched groups according to pretest performance and interest questionnaire ratings. The groups were divided into a 2 (KR by no KR) by 2 (do best goal by hard goal) design. The addition problems were arranged in a file box, and goals were made by placing a marker card in the box. The hard goal subject was told to try to reach the marker card. KR was administered at the end of

each trial when the subjects were told the number of problems they had gotten correct. The results indicated clearly that hard goal subjects out performed the do best subjects. There was no significant KR effect.

On the next study in this sequence, Locke and Bryan (1969a) replaced the do best condition with a specific easy goal condition. They also brought intrinsic goal setting under better control by presenting problems one by one on a spool instead of using the file box. Subjects were informed of their progress relative to the goal by control lights which were lit between trials. The KR subjects were informed of the number of problems attempted on each trial. Subjects who were given hard goal instructions significantly out performed the easy goal subjects while KR had no significant impact.

Another study was designed to consider the influence of KR and GS on reaction time (Locke, 1968b). One group of subjects was given no KR. A second group received green light KR every time they beat a standard and red light KR if they did not. The third group was given actual raw score information delivered by intercom. The green light - red light groups were divided according to the standard they had to beat to get the green light. One group had to beat their immediately previous score, another had to beat their worst score, while a third had to beat their best score. The results demonstrated that

the raw score KR group out performed the no KR group, the "immediately previous" standard group, and the "worst" standard group. The author states that interpreting the results as indicating a KR vs. no KR difference would be misleading. They found that the rate of improvement of the various standard groups was significantly different. The "best" and the "immediately previous" groups improved at a greater rate than the "worst" group. Since each group was receiving the same amount of information on each trial, the differences could only be explained by differences in goal setting. On a postexperimental questionnaire it was determined that subjects in the "raw score" KR group were assuming specific hard goals.

The KR-GS interaction was further studied when Locke and Bryan (1969b) studied two KR vs. KR + GS situations by independently varying KR and GS on each of two dependent variables. In a pilot study subjects worked simple arithmetic problems with knowledge of score following each trial. On half of the trials, they were to minimize their errors. In both instances the instructions significantly influenced the performance. A full experiment was then carried out when thirty subjects drove a Rambler station wagon around 3.4 mile course three times. Five dependent variables were measured: accelerator reversals (when the accelerator was depressed and let up); steering reversals (when the steering direction was changed); brake

applications; running time (length of trip minus stops); and trip time. All subjects took a turn around the course as a pretest, and then were given KR on each dimension measured. A counterbalanced design was used on the second and third trials. Half the subjects were given accelerator reversal goals first and steering reversal goals second. The goals were reversed for the second experimental group. The control group was given no KR and simply drove the course three times. The findings demonstrated the dramatic influence of goals or intentions on complex motor tasks. The groups scored according to the goals which they set. Each of these studies underlines the importance of GS and questions the role of KR.

Sauer (1970) found contradictory data. Sauer hypothesized that giving subjects rate of performance feedback would result in their setting higher rate of performance goals, while giving subjects accuracy feedback would result in higher accuracy goals. He further hypothesized that performance following positive feedback would differ from performance following negative feedback. The influence of goals was also investigated. Subjects given goals to perform faster and more accurately were expected to complete more of the task than subjects who were given no goals. Sauer also predicted that goal setting would reduce the variability between the KR and no KR conditions. To test these

hypotheses, 120 college students worked simple addition problems and recorded the answers on key punch machines. Each student was isolated in a small sound proof room and given predetermined KR. The basic experimental design was a 3 (rate of performance KR x accuracy KR x no KR) by 2 (goal x no goal) factorial. The dependent variables were the number of problems attempted and the number correct. Improvement in both of these measures was analyzed. A questionnaire was also completed at the end of the experiemtn to ascertain actual goal levels.

Although Sauer's results were largely nonsignificant, some interesting findings did appear. He found that while feedback was related to performance, goal setting was not. Specifically, subjects receiving rate of performance feedback added faster and less accurately than those receiving percent correct feedback. While KR variables were influencing performance, goal setting instructions had only an insignificant impact. A further negative finding was that no differences in intention were found to be related to feedback. The strongest predictor of tast performance was past performance. A second order relationship between feedback and goal setting was found when it was discovered that the influence of informative feedback on within cell variability was significantly greater when no goal had been set.

Sauer interprets his findings as demonstrating the importance of feedback while questioning the position of conscious goal setting. Only small differences appeared on a postexperimental questionnaire which was supposed to measure intentions. Negative findings on questionnaires are difficult to interpret for it could simply be that the instrument had very little power. Sauer inverted Locke's position when he concluded that goal setting did not influence performance but rather resulted from performance.

A second paper which has questioned Locke's conclusions was written by Cummings, Schwab, and Rosen (1971). These authors investigated the influence of performance and KR on goal setting. College students were again the subjects, and they worked on an addition task. The experiment involved two sessions separated by a short break. Subjects were given a pretest and then divided into four groups. Group 1 was given complete accurate KR; group 2 was given incomplete but accurate KR; the third group was given no KR; while the fourth group was given erroneously low KR. After KR was distributed all subjects were asked to set a goal for the second session. It was hypothesized that performance would influence goal setting. It was also suggested that after performance effects had been partialed out, the highest goals would be set by group 1, then group 2, group 3 and lastly group 4. Cummings et al, using a multiple regression

technique found that performance did influence goal setting. They also found that the correct KR group set significantly higher goals than the no KR group. The differences between the other KR groups was small. It was conlcuded that KR and previous performance clearly influenced goal level. It was suggested that these relationships be considered when goal setting, knowledge of results and performance are studied. Taken together these various results point to the interrelatedness of these variables. While definitive statements of cause and effect are premature, further research is warranted.

Socioeconomic Status, Sex, and Anxiety Level: Areas of Additional Interest. KR and GS are certainly not the only important variables which influence test performance. Three other potentially important factors are socioeconomic status (SES), sex, and anxiety level. The influence of SES and KR on Wechsler Intelligence Scale for Children scores was studied by Sweet and Ringness (1971). They varied race and referral status as will as KR and SES, but they studied only elementary school males. The results indicated that while lower class white subjects performed better under KR plus reinforcement conditions than under no KR, there were no significant differences for lower class blacks, middle class whites, nor the unreferred control group. An informal postexperimental question session revealed that KR subjects were engaging in goal

setting. The authors admitted that goal setting was an uncontrolled confounding variable. They concluded that motivational factors ought to be considered separately in studies which involve IQ testing.

Hokanson, Willers, and Koropsak (1968) found sex differences when subjects reacted to being electrically shocked by experimental confederates who were disguised as peers. Vascular recovery time was the dependent variable. Females calmed down significantly faster when they rewarded a "peer" who had just shocked them. Males, on the other hand, recovered more quickly when they returned a shock with a shock. This finding implies that there may be important differences between how males and females respond to stress in a testing situation.

Level of test anxiety, sex, and KR were varied by Campeau (1968). Fifth graders answered programmed instruction lessons dealing with the relationship between the earth and the sun. The results indicated that there was an anxiety by KR interaction for girls but not for boys. High anxiety girls scored better than low anxiety girls in the KR condition. Relative performance between high and low anxiety girls was reversed when no KR was provided. KR in this programmed instruction was viewed as reducing the test-like qualities and therefore reducing stress.

In general, the literature reviewed points out that knowledge of results and goal setting relate to testing and to each other. This is an area of controversy, of some confusion, but of considerable interest. The problem investigated in the present research was the separate and combined influences of KR and GS on test scores. SES, sex, and level of test anxiety were of tangential interest and were studied in after the fact analyses.

Statement of the Problem. Immediate knowledge of results and goal setting have a positive influence on skill acquisition. Educational and psychological literature have acclaimed feedback and goal setting as useful techniques. Students and instructors alike have reported that IKR and GS increase motivation in an instructional situation. Their impact on test performance, however, remains unproven. This is particularly so when different socioeconomic groups are compared.

Recent work with college students has suggested that immediate knowledge of results may detract from test performance. When the test counts for a grade, students have reported that knowing that they will be given results after each response makes them fell uneasy. Tests do count, and do put children

under a certain amount of pressure. How the children respond to this situation is a matter of theoretical and practical importance. The influence of IKR, GS, anxiety, SES, and sex on junior high school test performance is the subject of the present study. A postexperimental questionnaire will investigate the influence of the same independent variables on intention (goal setting), and will provide a rough measure of relative nervousness.

The .05 alpha level will be used to test the following experimental hypotheses. It is predicted that:

1. Subjects who receive incentive IKR score lower on an academic test than subjects who receive no IKR.

2. Subjects who receive explicit goal setting instructions score higher on an academic test than subjects who receive "do best" instructions.

3. Male and female subjects score differently on the Sequential Test of Educational Progress, Science Section.

4. The type of goal setting instructions that a subject receives (specific vs. "do best") interacts with immediate knowledge of results and sex when test score is the dependent variable.

5. Goal setting instructions and IKR influence grade estimates, as measured on a postexperimental questionnaire.

Subjects given specific goal instructions and those given IKR state higher goals than "do best" goal subjects.

6. IKR subjects rate the relative difficulty of the experimental portion of the test (part 2) higher than the no IKR subjects do.

7. IKR subjects rate their relative nervousness on the experimental part of the test (part 2) higher than the no IKR subjects do.

8. Type of goal setting instructions influence the goals that the subjects report on a postexperimental questionnaire. Goal setting instructions will result in students working toward a specific target goal.

9. Middle class subjects score higher than low class subjects on a multiple choice academic test that counts.

10. Level of socioeconomic status (middle vs. low) interacts with IKR vs. no IKR.

11. Subjects who rate themselves as having relatively high test anxiety will make more errors on the test and will be influenced differently by IKR than subjects who rate themselves as having relatively low test anxiety.

12. Level of socioeconomic status and test anxiety influences grade estimates, as measured on a postexperimental questionnaire.

13. A significant correlation exists between socio-economic status and test anxiety.

14. A significant predictor of multiple choice academic test performance is past performance on a similiar test.

15. Pretest performance on a multiple choice academic test influences estimation of part 2 score (goal setting), as measured on a postexperimental questionnaire. Subjects who scored high on a pretest set higher goals than subjects who scored low on a pretest.

16. Grade estimates (goal setting), as measured on a postexperimental questionnaire reflect test scores. Subjects who have set high grade estimates score relatively high on the test, while subjects who set low grade estimates score relatively poorly on the test.

Chapter 2

METHOD

Subjects. The subjects were the entire seventh grade from the Orange County, Virginia, Intermediate School. Orange County is located in central Virginia and is primarily rural. There was a total subject population of 304 pupils. Subjects were placed in treatment cells such that an equal number of males and females were in each condition. The cells were KR vs. no KR and GS vs. "do your best." Aside from matching for sex, students were assigned to cells randomly. Subjects who missed one or both of the testing sessions were dropped from the study.

Instruments. A test consisting of questions from Form 4A (fourth to sixth grade) of the science section of the Sequential Test of Educational Progress (STEP) was the major dependent variable. The STEP test was selected because it was appropriate for the new seventh graders serving as subjects. The test was part of an achievement battery which could provide information as to what students know at the beginning of the school year. The test has the ability to discriminate between subjects and it has proven reliability (ETS, 1956, 1969). A

complete technical manual was available with the test. The manual demonstrated that the test has been modernized and revalidated. The science test was given in two parts on adjoining days. Part 1 included the first thirty items from Form 4A, Series 1 of the STEP test. Part 2 was the first forty items on Form 4A, Series 2 of the STEP test. The questions on part 2 were selected so there would be no possibility of IKR on one item providing information helpful on another item.

Three additional independent instruments were used to analyze a posteriori data. Students were questioned regarding their sex and about their parents' occupations (see Appendix 1 for questionnaire). The occupation information was then used to divide the students into SES quartiles according to Roe's scale (1956). Students were also asked six questions from the "Test Anxiety Scale for Children" (Sarason, Lighthall, Davidson, Waite, Ruebush, 1960; see Appendix 2). The items were selected because they dealt specifically with test situations, and were the kinds of questions that accounted for the most variance on the test anxiety scale (Dunn, 1965). A second dependent measure was a postexperimental questionnaire (note Appendix 3) which attempted to ascertain the students' goal intentions (question 3) and a measure of nervousness (question 8).

Procedure. The testing took place during regular science periods at the beginning of the school year. The tests were administered and monitored by the regular science teachers and the experimenter who had been introduced as an aide.

The test was introduced as an examination that counted for section placement as well as for a grade. The test was divided into two parts of comparable difficulty. Both parts were answered on specially treated response cards which were available commercially from Van Valkenburgh, Nooger, and Neville, Inc., New York, New York. Subjects indicated a response by erasing a mark associated with the alternative they selected. On part 1 all subjects used response cards from Lot Z11 Second Series. The graphite marks cover a meaningless assortment of letters, "T", "E", "H", and "L."

The teacher began by reading the following instructions to the children:

> "This is a science test. The test is different from those which you took last year so please pay close attention to the instructions. There are thirty questions on the test. Read each question carefully and then select the answer that you think is best. When you have made your choice, look over to the little card that came with your test. Find the question and the letter of the choice that you think is correct. Now erase the mark next to the letter which you have chosen. You must erase firmly to remove the mark, but do not erase so hard that you hurt the letter under the mark. (Then the teacher will demonstrate.) Be sure that you take your time because once

> you have erased a mark you can not change your mind. If you erase two marks, your answer will be counted wrong. Your score will be based on the total number you get right so you are free to guess if you are not sure of an answer. You will have the whole period to complete this test. There is to be no talking during the test. If you have any questions, raise your hand and someone will come and help you."

Each student was given an identical test packet which included a cover sheet, test book, and response card. The cover sheet had the following instructions on it:

> "This is a science test. Read each question carefully and select the best answer. When you have made your choice, record it by erasing the mark next to the letter. After you erase a mark, you will notice that a letter will appear. Different letters will appear at different times. They do not mean anything. No talking is permitted. If you have any questions, please raise your hand and someone will come to help you."

On part 2 the independent variables, IKR and GS, were used. Half of the subjects were provided IKR via the response cards (Lot Z11b). The IKR directions indicated that uncovering a "T" meant a correct response while uncovering another letter meant incorrect. The instructions read:

> "This is a science test. Read each question carefully and select the best answer. When you have made your choice, record it by erasing the mark next to the letter. You will be able to tell whether you are right or wrong after you erase a mark. A letter 'T' will show if your answer is right. If you are wrong, another letter will show. Remember if you erase a mark and find a 'T', it means you are correct. Any other letter means that you are wrong. No talking is permitted. If you have any questions, please raise your hand and someone will come to help you."

The no IKR subjects received written instructions identical to those on part 1. The arrangement of the hidden letters on their response cards (Lot Z11) was different from that on the IKR subjects' cards.

The second independent variable was introduced when half of the subjects received goal setting instructions. All subjects received part 2 with their names written on it to allow for individual instructions which were printed on the cover sheet of the test packet.

The GS students' instructions read:

> "You took part 1 yesterday. That test was graded. I have an idea of how well you can do on a test like this. To be successful, you should aim to get _____ (a number 20% greater than the subject got right on part 1) questions right on this section of the test. To get _____ questions right, you have to get about _____ (1/8 of his goal rounded up to the nearest whole number) out of each five questions correct."

The "try your best" GS students' instructions read:

> "You took part 1 yesterday. That test was graded. I have an idea of how well you can do on a test like this. To be successful, you should try your best on this section of the test."

Experimental Design. The study used a 2 (IKR vs. no IKR) x 2 (goal setting instructions vs. no goal setting instructions) x 2 (male vs. female) design. This permitted three main effects and four interactions to be measured. The score on part 1 was taken into account through the use of an analysis of covariance

technique.

The data was then redivided and reanalyzed. The subjects were classfied by their sex, anxiety level, and socioeconomic status (SES), as well as by their IKR and GS condition. Four 2 x 2 analyses of variance were performed to test whether IKR and GS influenced grade estimates, and whether SES and test anxiety influenced test score and grade estimates.

A Chi Square techniques was used to delineate the impact of IKR on perceived item difficulty and relative nervousness. Finally correlations were determined between SES and anxiety, test scores on parts 1 and 2, and between grade estimates and actual scores on parts 1 and 2.

Chapter 3

RESULTS

The major purpose of the analysis was to determine the influence of immediate knowledge of results (IKR) and goal setting instructions (GS) on test performance. Supplementary analyses were undertaken to test for the impact of sex, anxiety level, and socioeconomic status upon test score and upon grade expectations (goals). Four different techniques were used in order to fully evaluate the various hypotheses. These included 2 x 2 x 2 analysis of covariance, 2 x 2 analyses of variance, analyses by Chi Square and analyses by correlation.

Hypotheses 1, 2, 3, and 4:

1. Subjects who receive incentive IKR score lower on an academic test than subjects who receive no IKR.

2. Subjects who receive explicit goal setting instructions score higher on an academic test than subjects who receive "do best" instructions.

3. Male and female subjects score differently on the Sequential Test of Educational Progress, Science Section.

4. The type of goal setting instructions that a subject receives (specific vs. "do best") interacts with immediate knowledge of results and sex when test score is the dependent variable.

Table 1 shows the results relevant to these hypotheses. All cells were reduced randomly to arrive at equal cell frequencies (n = 33). A 2 x 2 x 2 (IKR x GS x sex) analysis of covariance failed to reveal any significant differences resulting from the influence of the three independent variables. The means for the IKR vs. no IKR condition were in the direction opposite that predicted.

Hypothesis 5:

5. Goal setting instructions and IKR influence grade estimates, as measured on a postexperimental questionnaire. Subjects given specific goal instructions and those given IKR state higher goals than "do best" goal subjects.

Cell frequencies were randomly reduced to obtain equal numbers of subjects in each cell (n = 72). The hypothesis was tested by a 2 x 2 analysis of variance. The data, as listed in Table 2, failed to support the hypotheses.

Hypothesis 6:

6. IKR subjects rate the relative difficulty of the experimental portion of the test (part 2) higher than the no IKR subjects do.

The hypothesis was tested using a Chi Square technique (note Table 3). The relative difficulty ratings (question 6 on the questionnaire) of the two parts of the test for the

Table 1

Immediate Knowledge of Results by Goal Setting Instructions by Sex, where Test Score on Part Two was the Dependent Variable

Adjusted Means

	GS		No GS	
	male	female	male	female
IKR	14.04	16.04	15.45	15.91
No IKR	14.25	16.19	15.73	13.43

2 x 2 x 2 Analysis of Covariance

Source	df	MS	F	
IKR (A)	1	65.3178	3.2019	1.
GS (B)	1	33.2546	1.6301	2.
Sex (C)	1	75.5493	3.7035	3.
A x B	1	45.3715	2.2241	3.
A x C	1	1.2417	0.0608	3.
B x C	1	20.6434	1.0119	3.
A x B x C	1	3.0217	0.1481	3.
Error	255	20.3994		

1. $F_{.95}(1,255) = 2.72$ for one tail test (however, means are in opposite of the predicted direction)
2. $F_{.95}(1,255) = 2.72$ for one tail test
3. $F_{.95}(1,255) = 3.88$ for two tail test

two groups, IKR and no IKR, were compared. Although the Chi Square was significant beyond the .05 level, inspection of the table revealed that the difference was not in the predicted direction. IKR subjects rated the experimental section (part 2) of the test as relatively easy while the no IKR group rated it as relatively difficult.

Hypothesis 7:

7. IKR subjects rate their relative nervousness on the experimental part of the test (part 2) higher than the no IKR subjects do.

The Chi Square method was again used (Table 4) to compare the questionnaire data (question 8). Although the X^2 was not significant, the trend was in the direction of IKR reducing nervousness. This trend was opposite that which was predicted.

Hypothesis 8:

8. Type of goal setting instructions influence the goals that the subjects report on a postexperimental questionnaire. Goal setting instructions will result in students working toward a specific target goal.

The data failed to support this hypothesis: 268 subjects responded to the question by saying that they tried their best, 4 said that they did not try, and 32 said they aimed for a specific goal. Seventeen of these 32 had received goal setting instructions; 15 had received "try best" instructions.

Table 2

Immediate Knowledge of Results (IKR) and Goal Setting Instructions (GS) where Grade Estimate was the Dependent Variable

Means

	GS	No GS
IKR	29.7	25.6
No IKR	28.6	26.3

Analysis of Variance

Source	df	MS	F	
GS (A)	1	84.5000	1.2413	1.
KR (B)	1	30.0556	0.2946	1.
A x B	1	8.6806	0.1275	2,
Error	284	68.0730		

1. $F_{.95}(1,284) = 2.72$ for one tail test

2. $F_{.95}(1,284) = 3.87$ for two tail test

Table 3

IKR and Relative Difficulty of the Two Tests:

Observed and Expected Frequencies

		Part 1 relatively easy	neither	Part 1 relatively hard	Total
IKR	observed	36	71	33	140
	expected	44.4	68.8	27.3	
No IKR	observed	50	63	20	133
	expected	42.0	65.7	25.9	
	Total	86	134	53	273

df = 2 $p < .05^{*} = 5.99$

$X^2 = 6.52^{*}$ (frequencies in opposite direction from prediction)

Table 4

IKR and Relative Nervousness:

Observed and Expected Frequencies

		more nervous yesterday	neither	more nervous today	Total
IKR	observed	41	78	33	152
	expected	34.3	85.0	32.3	
No IKR	observed	27	91	31	149
	expected	33.6	83.1	31.6	
	Total	68	169	64	301

df = 2 $p < .05^{*} = 5.99$

$X^2 = 3.96$

Hypotheses 9 and 10:

9. Middle class subjects score higher than low class subjects on a multiple choice academic test that counts.

10. Level of socioeconomic status (middle vs. low) interacts with IKR vs. no IKR.

Cell frequencies were randomly reduced to obtain equal numbers of subjects in each cell (n = 28). These hypotheses were tested by using a 2 x 2 analysis of variance. The results, as listed in Table 5, demonstrate that high SES is associated with relatively high test score, but no significant effect was found for IKR or interaction.

Hypothesis 11:

11. Subjects who rate themselves as having relatively high test anxiety will make more errors on the test and will be influenced differently by IKR than subjects who rate themselves as having relatively low test anxiety.

Again cell frequencies were randomly reduced to obtain equal numbers of subjects in each cell (n = 41). A 2 x 2 analysis of variance was carried out. Strong support was found for the impact of test anxiety on test performance; those who rated themselves as relatively anxious made more errors than those who rated themselves as less anxious.

Table 5

Socioeconomic Status (SES) and Immediate Knowledge of Results (IKR) Where Error Score on Part 2 was the Dependent Variable

Means

	IKR	No IKR
High SES	10.8	12.4
Low SES	17.4	17.5

Analysis of Variance

Source	df	MS	F	
IKR (A)	1	16.5089	0.4117	2.
SES (B)	1	952.7232	23.8127	1.**
A x B	1	16.5089	0.4117	2.
Error	108	40.0929		

1.** $F^{.99}_{(1,108)} = 2.75$ for one tail test

No evidence of an influence by IKR and no significant interaction was discovered. These results are listed in Table 6.

Hypothesis 12:

12. Level of socioeconomic status and test anxiety influences grade estimates, as measured on a postexperimental questionnaire.

A median split was used to divide the students into middle and low SES groups. Cell frequencies were randomly reduced to obtain equal cell size (n = 21). A 2 x 2 analysis of variance (SES and test anxiety) was carried out to test the hypothesis. The results are listed in Table 7. The data failed to significantly support the hypothesis.

Hypothesis 13:

13. A significant correlation exists between socioeconomic status and test anxiety.

A high number of the abbreviated Sarason scale meant a relatively high anxious person. In a like manner, a high number on the SES scale meant a relatively low SES occupation. It was therfore predicted that a positive and significant correlation would exist. A correlation of 0.17 was found: it was significantly different from zero, $p < .005$.

Table 6

Test Anxiety Level and Immediate Knowledge of Results (IKR) where Part 2 Error Score was the Dependent Variable.

Means

	High Anxiety	Low Anxiety
IKR	18.4	11.6
No IKR	18.2	13.6

Analysis of Variance

Source	df	MS	F	
Anxiety (A)	1	1080.7378	27.5099	1.**
IKR (B)	1	32.4939	0.8271	2.
A x B	1	10.2500	0.2609	2.
Error	160	39.2853		

1.** $F_{.95}(1,160) = 3.90$ for one tail test

Table 7

Test Anxiety and Socioeconomic Status (SES) where Grade Estimate on Part 2 was the Dependent Variable

Means

	High SES	Low SES
High Test Anxiety	27.1	30.9
Low Test Anxiety	26.1	29.4

Analysis of Variance

Source	df	MS	F	
Anxiety (A)	1	33.4405	0.3999	2.
SES (B)	1	264.2976	3.1612	2.
A x B	1	1.4405	0.0172	2.
Error	80	83.6047		

$F_{.95}(1,80) = 3.98$ for a two tail test

Hypothesis 14:

14. A significant predictor of multiple choice academic test performance is past performance on a similiar test.

The data strongly supports the expectation that pretest (part 1) performance correlated with criterion test (part 2) performance, $r = 0.80$ for 304 subjects, $p < .0005$.

Hypothesis 15:

15. Pretest performance on a multiple choice academic test influences estimation of part 2 score (goal setting), as measured on a postexperimental questionnaire. Subjects who scored high on a pretest set higher goals than subjects who scored low on a pretest.

A Pearson coefficient of correlation was taken between part 1 scores and part 2 score estimates. A significant relationship existed: $r = .32$, $p < .0005$.

Hypothesis 16:

16. Grade estimates (goal setting), as measured on a postexperimental questionnaire reflect test scores. Subjects who have set high grade estimates score relatively well on the test, while subjects who set low grade estimates score relatively poorly on the test.

This hypothesis was tested with a Pearson correlation

coefficient. A significant correlation was found $r = 0.42$, $p < .0005$. In order to more fully investigate this strong relationship, the subjects were broken into four cells (IKR vs. no IKR and GS vs. no GS) and correlations were measured for each cell. The correlations are listed in Table 8.

The differences among the accuracy of the estimates of the four groups were tested using the Fisher r to z method, no significant differences were found.

Table 8

Table of Correlations between Score on Part 2 of the Test and Subject's Score Estimates as Measured by a Postexperimental Questionnaire

		Immediate Knowledge of Results	
		IKR	no IKR
Goal Setting Instructions	GS	.4911	.3069
	No GS	.5604	.3343

r to z Transformation

GS - IKR vs. GS - no IKR: .4911 vs. .3069 z = 1.31

No GS - IKR vs. no GS - no IKR: .5604 vs. .3343 z = 1.71

GS - no IKR vs. no GS - IKR: .3069 vs. .5604 z = 1.91

t .95 (∞) = 1.96

Chapter 4

DISCUSSION

The results of this investigation failed to demonstrate a significant influence of the independent variables (IKR and GS instructions) on the dependent variable (test score). These findings must be viewed in light of the results of hypotheses eight and sixteen. These hypotheses questioned the actual functionality of the two independent variables. Subjects were asked about their intentions during the test to determine whether the goal setting instructions actually were adopted (hypothesis 8). No support was found for this hypothesis. Regardless of the instructions, subjects overwhelmingly stated that they "tried their best."

The functionality of IKR was measured when test score estimates were taken after the examination to delineate the influence of IKR on the subject's ability to guess his test scores (hypothesis 16 as shown in Table 8). It was hypothesized that a subject who had been told whether each response was correct or not would be able to more accurately estimate his success on the test than a subject who had been deprived of feedback. However, no support was found for this hypothesis. The subject's accuracy regarding how well he had done on the test was about the same regardless of the feedback

condition. Adding extrinsic item by item knowledge of results failed to significantly increase the correlation between subject's estimated test score and subject's actual score. Apparently these subjects did not use the IKR from the response cards (obtained commercially from Van Valhenburg, Nooger, and Neville, Inc.) to improve their idea of how well they were doing on the test as a whole. The data demonstrated that the IKR failed to be functional.

A variety of reasons may expalin the possible lack of functionality. Perhaps the students were not interested in whether their answers were correct or not. It was conceivable that individual test items may have been so factually oriented and differentially easy or hard that the students knew whether they were right of wrong. If that had been the case, providing IKR would have served no useful purpose. If more ambiguous analogy or comprehension items had been used, they may have offered less intrinsic IKR and thereby provided extrinsic IKR with an opportunity to have an impact. It was also possible that the students did not understand the directions and were never aware that they were receiving IKR. Since experimental directions were read by each student at his (or her) seat, it was possible that the students were unable to understand the procedure. A final explanation may be that the subjects

received the IKR, but failed to accumulate it to improve their idea of how well they had done on the examination. If the students' perceived performance was below their expectancy level for academic tests, perhaps they ignored the IKR to reduce cognitive dissonance.

The correlation between previous test score and criterion score was .80, while the correlation between estimated scores and actual scores for IKR subjects was only .49 and .56 for the GS and no GS groups respectively. Thus one could have gotten more information about how students had scored on the test by looking at past scores than by asking the students themselves. This occurred in spite of the fact that the students had been given item by item IKR. It was conceivable that the subjects had a mental set relating to how well they did on tests, and that telling them whether they got items correct or not did not alter that set. This suggested an additional analysis which demonstrated that students' grade estimates correlated only slightly lower with their previous day test score, $r = .32$ than with their criterion test score.

The findings relating to hypotheses 1, 2, and 3 failed to demonstrate that there was a significant influence of the GS instructions and IKR on test score. If any difference of psychological importance existed on this task with these

students, 304 subjects would have been expected to demonstrate that difference.

It was noteworthy that contrary to the hypotheses, the trend in the data suggested that the IKR students did somewhat better than the no IKR students. This was the case in spite of the fact that all questions were examined to be sure that knowledge of results on one question would not assist the student in answering another question. The conclusions regarding the negative influence of IKR for college students (Bierbaum, 1968; Strang and Rust, 1973) did not appear to apply in the current experiment. In fact the evidence from the postexperimental questionnaire indicated that the IKR influenced the subjects by making them regard the experimental section of the test (part 2) as relatively easy.

Some important limitations were evident in the current investigation. The results could not be made to apply to other goal setting instructions or to other modes of IKR. Similiarly the demographic characteristics of the area in question limit the generalizability of the study.

Test anxiety was demonstrated to be an intervening variable of considerable importance (hypothesis 13). The subjects' measured anxiety scores were related to their test performance (hypothesis 11). While there was no evidence as to

a cause and effect relationship, the data did support the importance of the construct, test anxiety, as accounting for variability among test scores. The strength of test anxiety as a construct was underlined by the fact that only a six item questionnaire, an abbreviated Sarason Scale, was used to divide the students into high and low anxiety classifications.

Other confounding factors appeared to work in conjunction with test anxiety. These included the time of year (the first week of school) and the testing condition (a nearly silent room). The students were told that the test was to be used for placement and that they were to remain silent. The children had the entire period to complete the test. Although this was more time than recommended in the test instructions, few of the students finished before time was called. Three proctors maintained a strict vigil over each group of approximately 50 students. The combination of factors (beginning of year + rural students + placement instructions + strict silence + time pressure) may have produced a psycho-situational climate so highly geared to motivation that IKR and GS factors were partially masked. This conclusion was supported by anecdotal comments from the proctors: the students appeared to be intently concentrating on the test. Of course, these are the conditions frequently found in realistic testing situations,

and it is of greater interest to determine the influence (or lack of influence) of IKR under such realistic conditions rather under construed labratory conditions.

Sex was not found to be a significant factor on test performance. However, there was a trend indicating that the boys made fewer errors on the science test than the girls. This tendency may have been due to the particular sex related characteristics of the subjects. Perhaps the boys were more interested in the scientific questions, and more familiar with the terms used.

The fact that socioeconomic status (SES) was found to relate to test score was not surprising (Rucinski, 1968); children from middle class families consistently scored better than their lower class peers. It was also hypothesized that SES and test anxiety ratings were related (hypothesis 12). The resulting significant correlation 0.17 ($p < .005$) was of questionable psychological importance. Less than three percent of the variance of the test anxiety scores was accounted for by the subjects' social class.

The lack of the apparent functionality of the IKR and GS treatments was particularly interesting. It suggested that any further experiments in this area must consider the actual impact of the independent variables. Further, past inconsis-

tencies in studies in this area may have resulted from a failure to deal with this issue. Another question which was raised related to the possible power of the subject's mental set and how item by item feedback did little to change it. Further research might deal with the relationship between pretest score estimates, actual scores, IKR conditions, and post test estimates.

References

Ammons, R. B. Effects of knowledge of performance: a survey and tentative theoretical formulation. Journal of General Psychology, 1956, 54, 279-299.

Anderson, H. M., & Brandt, H. F. Study of motivation involving self announced goals of fifth grade children and the concept of level of aspiration. Journal of Social Psychology, 1939, 10, 209-232.

Anderson, R. C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-370.

Anderson, R. C., Kulhavy, R. W., & Andre, T. Feedback procedures in programmed instruction. Journal of Educational Psychology, 1971, 62, No. 2, 148-156.

Angell, G. W. The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. Journal of Educational Research, 1949, 42, 391-394.

Annett, J. Feedback and Human Behavior: The Effects of Knowledge of Results, Incentives and Reinforcement on Learning and Performance. Baltimore: Penguin Books, 1969.

Atkinson, J. W. Motivational determinants of risk-taking behavior. Psychological Review, 1957, 64, 359-372.

Atkinson, J. W., & Feather, N. T. A Theory of Achievement Motivation. New York: Wiley, 1966.

Battle, E. S. Motivational determinants of academic competence. Journal of Personality and Social Psychology, 1966, 4, 634-642.

Bayton, J. A. Performance as a function of expressed and non-expressed levels of aspiration. American Psychologist, 1948, 3, 274 (Abstract).

Berglund, G. The effect of knowledge of results on retention. Psychology in the Schools, 1969, 6(4), 420-421.

Berman, M. L. Shaping performance on programmed material. Psychological Reports, 1967, 21, 29-32.

Bierbaum, W. B. Immediate knowledge of performance on multiple choice tests, Journal of Programmed Instruction, 1965, 3, 19-23.

Bottenberg, R. A., & Ward , J. H. Applied Multiple Linear Regression. Technical Report PRL - TDL - 63 - 6, Aerospace Medical Division, Lackland Air Force Base, Texas, 1963.

Brackbill, Y., Wagner, J. E., & Wilson, D. Feedback delay and the teaching machine. Psychology in the Schools, 1964, 1, 148-150.

Brown, F. J. Knowledge of results as an incentive in school room practice. Journal of Educational Psychology, 1932, 23, 532-552.

Bryan, J. F., & Locke, E. A. Parkinson's law as a goal-setting phenomenon. Organizational Behavior and Human Performance, 1967, 2, 258-275. (a)

Bryan, J. F. & Locke, E. A. Goal setting as a means of increasing motivation. Journal of Applied Psychology, 1967, 51, 274-277. (b)

Bryant, B. & Anisfeld, M. Feedback versus no-feedback in testing children's knowledge of English pluralization rules. Journal of Experimental Child Psychology, 1969, 8, 250-255.

Campeau, P. L. Test anxiety and feedback in programmed instruction. Journal of Educational Psychology, 1968, 59(3), 159-163.

Clark, W. C., & Greenberg, D. B. The effect of stress, knowledge of results, and proactive inhibition on verbal recognition memory (D') and response criterion (LX). Journal of Personality and Social Psychology, 1971, 17, 42-47.

Clifford, M. M. Motivational effects of competition and goal setting in reward and non-reward conditions. The Journal of Experimental Education, 1971, 39(3), 11-16

Cofer, C. N., & Appley, M. H. Motivation: Theory and Research. New York: John Wiley and Sons, 1964.

Cummings, L. L., Schwab, D. P., & Rosen, M. Performance and knowledge of results as determinants of goal setting. Journal of Applied Psychology, 1971, 55(6), 526-530.

Dahle, A. J., & Daly, D. A. The influence of verbal feedback on auditory discrimination test performance of mentally retarded children. American Journal of Mental Deficiency, 1972, 76, 586-590.

Dey, M. K., & Kaur, G. Facilitation of performance by experimentally induced ego motivation. Journal of General Psychology, 1965, 73, 237-247.

Dowell, E. C. Gurney, P. J., & Norris, K. H. An evaluation of trainer-testers. Training Analysis and Development Directorate, Headquarters Technical Training Air Force, Gulf Port, Mississippi, 1956.

Dunn, J. A. Stability of the factor structure of the test anxiety scale for children across age and sex groups. Journal of Consultating Psychology, 1965, 29, 187.

Dunn, J. A. The approach-avoidance model for analysis of school anxiety. Journal of Educational Psychology, 1968, 59, 388-394.

Eason, R. G. Relation between tension level skill and performance efficiency on a perceptual motor task. Perceptual and Motor Skills, 1963, 16, 297-317.

Educational Testing Service. Sequential Test of Educational Progress. Princeton, New Jersey: Cooperative Test Division, 1956 and 1969.

Festinger, L. Wish, expectation, and group standards as factors influencing level of aspiration. Journal of Abnormal and Social Psychology, 1942, 37, 184-200.

Flook, A. J. M., & Saggar, U. Academic performance with and without knowledge of scores on tests of intelligence, aptitude, and personality. Journal of Educational Psychology, 1968, 59, 395-401.

Frank, J. D., Individual differences in certain aspects of the level of aspiration. American Journal of Psychology, 1935, 47, 119-128.

Frase, L. T. Learning from prose material: length of passage, knowledge of results, and position of questions. Journal of Educational Psychology, 1967, 58, 266-272.

Fryer, F. W. An Evaluation of Level of Aspiration as a Training Procedure. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.

Gaa, J. P. The use of individual goal setting conferences as a motivational technique. ERIC, University of North Carolina, 1970, ERIC 053-816.

Galdieri, A. A., Barcikowski, R. S., & Witmer, J. M. The effect of verbal approval upon the performance of middle- and lower-class children on the WISC. Psychology in the Schools, 1972, 9(4), 404-408.

Gaudet, I. J. & Moon, W. H . Effects of differential feedback on opinions of and responses to intelligence tests. Journal of Consulting and Clinical Psychology, 1970, 35 (3), 376-380.

Gilman, D. A. The effect of feedback on learners' certainty of response and attitude toward instruction in a computer. Journal of Research in Science Teaching, 1969, 6, 171-184.

Gray, R. T. An evaluation of the effect of an immediate feedback device used with a typical college classroom. ERIC, San Diego State College, 1968. ERIC 015-658.

Gutherie, J. T. Motivational effects of feedback in reading. ERIC, Johns Hopkins University, 1971. ERIC 042-590.

Heald, H . M. The effects of immediate knowledge of results and correction of errors and test anxiety upon test performance. Dissertation Abstracts International, 1970, 31(5-A), 1621-1622.

Hertzman, M. & Festinger, L. Change in explicit goals in a level of aspiration experiment. Journal of Experimental Psychology, 1940, 27, 439-452.

Hokanson, J. D., Willers, K. R. & Koropsak, E. The modification of autonomic responses during aggressive interchange. Journal of Personality, 1968, 36, 386-404.

Holding, D. H. Principles of Training. New York: Macmillan, 1965.

Holland, J. G. & Skinner, B. F. The Analysis of Behavior. New York: McGraw-Hill, 1961.

Hurlock, E. B. An evaluation of certain incentives used in school work. Journal of Educational Psychology, 1925, 16, 145-159.

Kaess, W. & Zeaman, D. Positive and negative knowledge of results on Pressey-type punchboard. Journal of Experimental Psychology, 1960, 60, 12-17.

Karraker, R. J. Knowledge of results and incorrect recall of plausible multiple choice alternatives. Journal of Educational Psychology, 1967, 58, 11-14.

Kausler, D. H. Aspiration level as a determinant of performance. Journal of Personality, 1959, 27, 346-351.

Kausler, D. H. & Trapp, E. P. Achievement motivation and goal setting behavior on a learning task. Journal of Experimental Psychology, 1958, 55, 575-578.

Kirby, T. J. Practice in the case of school children. Teachers College Contributions, 1913, 58, 43.

Krumboltz, J. D. & Weisman, R. G. The effect of intermittent confirmation in programmed instruction. Journal of Educational Psychology, 1962, 53, 250-253.

Labov, W. Academic ignorance and black intelligence. The Atlantic Monthly, 1972, 229, 59-67.

Lee, S. J. A study of goal expectation behavior in the context of level of aspiration: an identification of empirical principles. Dissertation Abstracts, 1968, 28(7-B) 3051-3052.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. Level of aspiration. In Hunt, J. (Ed.) Personality and the Behavior Disorders. Vol. 1. New York: Ronald Press, 1944. Pp. 333-378.

Locke, E. A. The relationship of intentions to level of performance. Journal of Applied Psychology, 1966, 50, 60-66. (a)

Locke, E. A. A closer look at level of aspiration as a training procedure: a reanalysis of Fryer's data. Journal of Applied Psychology, 1966, 50, 417-420. (b)

Locke, E. A. The motivational effects of knowledge of results: knowledge or goal setting? Journal of Applied Psychology, 1967, 51, 324-329.

Locke, E. A. Toward a theory of task motivation and incentives. Organizational Behavior and H uman Performance, 1968, 3, 157-189. (a)

Locke, E. A. Effects of knowledge of results, feedback in relation to standards, and goals on reaction-time performances. American Journal of Psychology, 1968, 81, 566-574. (b)

Locke, E. A. & Bryan, J. F. The effects of goal setting, rule learning and knowledge of score on performance. American Journal of Psychology, 1966, 79, 451-457. (a)

Locke, E. A. & Bryan, J. F. Cognitive aspects of psychomotor performance: the effects of performance goals on level of performance. Journal of Applied Psychology, 1966, 50, 286-291. (b)

Locke, E. A. & Bryan, J. A. Performance goals as determinants of level of performance and boredom. Journal of Applied Psychology, 1967, 51, 120-130.

Locke, E. A., & Bryan, J. F. Grade goals as determinants of academic achievement. Journal of General Psychology, 1968, 79, 219-228 (a)

Locke, E. A. & Bryan, J. F. Goal setting as a determinant of the effect of knowledge of score on performance. American Journal of Psychology, 1968, 81, 398-406. (b)

Locke, E. A. & Bryan, J. F. Knowledge of score and goal level as determinants of work rate. Journal of Applied Psychology, 1969, 53, 59-65. (a)

Locke, E. A. & Bryan, J. F. The directing function of goals in task performance. Organizational Behavior and Human Performance, 1969, 4, 35-42. (b)

Locke, E. A., Bryan, J. B., & Kendall, L. M. Goals and intentions as mediators of the effects of monetary incentives on behavior. Journal of Applied Psychology, 1968, 52, 104-121.

Locke, E. A., Cartledge, N., Knerr, C. S., & Bryan, J. F. The Motivational Effects of Knowledge of Results. Washington, D. C.: American Institutes for Research, 1969.

Locke, E. A., Cartledge, N., & Koeppel, J. Motivational effects of knowledge of results: a goal setting phenomenon? Psychological Bulletin, 1968, 70, 474-485.

Lublin, S. C. Reinforcement schedules, scholastic aptitude, autonomy, and achievement in programmed instruction course. Journal of Educational Psychology, 1965, 56, 295-302.

Mace, C. A. Incentives: some experimental studies. Industrial Health Board (Great Britain) 1935, Report #72.

Manzer, G. W. The effect of knowledge of output on muscular work. Journal of Experimental Psychology, 1935, k8, 80-90.

Marso, R. N. Classroom testing procedure, test anxiety, and achievement. Journal of Experimental Education, 1970, 38, 54-58.

McClelland, D. C. The Achieving Society. Princeton, New Jersey: Vern Nostrand, 1961.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, F. L., The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

Means, R. S., & Means, G. H. Achievement as a function of the presence of prior information concerning aptitude. Journal of Educational Psychology, 1971, 62, 185-187.

Montor, K. Affect of using a self-scoring answer sheet on knowledge retention. Journal of Educational Research, 1970, 63(10), 435-437.

Morgan, C. L. & Morgan, L. U. Effects of immediate awareness of success and failure upon objective examination scores. Journal of Experimental Education, 1935, 5, 63-66.

Orne, M. T. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783.

Paige, D. D. Learning while testing. Journal of Educational Research, 1966, 59, 276-277.

Payne, R. B. & Hauty, G. T. Effect of psychological feedback upon work decrement. Journal of Experimental Psychology, 1955, 50, 343-351.

Pressey, S. L. Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. Journal of Psychology, 1950, 29, 417-447.

Pressey, S. L. Teaching machine and learning theory crisis. Journal of Applied Psychology. 1963, 47, 1-6.

Pressey, S. L. Re-program programming. Psychology in the Schools, 1967, 4 (3), 234-239.

Pressey, S. L. & Kinzer, J. R. Auto-elucidation without programming. Psychology in the Schools, 1964, 1 (4), 359-365.

Roe, A. The Psychology of Occupation. New York: John Wiley, 1956.

Rucinski, P. R. Motivation effect of two reinforcers upon lower- and middle-class fifth grade children. Journal of Educational Research, 1968, 61, 368-371.

Ryan, T. A. Intentional Behavior. New York: Ronald Press, 1970.

Ryan, T. A. Drives, tasks, and the initiation of behavior. American Journal of Psychology, 1958, 71, 74-93.

Sarason, S. B., Davidson, K., Lighthall, F., & Waite, R. A test anxiety scale for children. Child Development, 1958, 29, 105-113.

Sarason, S. B., Lighthall, F. F., Davidson, K. S., Waite, R. R., & Ruebush, B. K. Anxiety in Elementary School Children. New York: John Wiley, 1960.

Sassenrath, J. M. & Yonge, G. D. Delayed information feedback, feedback cues, retention set, and delayed retention. Journal of Educational Psychology, 1968, 59, 69-73.

Sauer, J. R. The relationship between informative feedback and goal setting. Dissertation Abstracts International, 1970, 30 (10-B), 4782.

Stedry, A. C. Budget Control and Cost Behavior, Englewood Cliffs, New Jersey: Prentice-Hall, 1960.

Strang, H. R. Task definition and immediate knowledge of results in monitoring performance. Perceptual and Motor Skills, 1972, 34, 825-826.

Strang, H . R. & Rust, J. O. The effects of immediate knowlege of results and task definition on multiple choice answering. Journal of Experimental Education, in press.

Sullivan, H. J., Baker, R. L., & Schultz, R. E. Effect of intrinsic and extrinsic reinforcement contingencies on learner performance. Journal of Education Psychology, 1967, 58 (3), 165-169.

References

Sweet, R. C. & Ringness, T. A. Variations in the intelligence test performance of referred boys of idffering racial and socioeconomic backgrounds as a function of feedback or monetary reinforcement. Journal of School Psychology, 1971, 4, 399-409.

Terman, L. M., & Merrill, M. A. Measuring Intelligence. Boston: Houghton Mifflin, 1937.

Thorndike, E. L. The Fundamentals of Learning. Teacher's College, Columbia University, 1932.

Travers, R. M. W., Van Wagenen, R. K., Haygood, D. H., & McCormick, M. Learning as a consequence of the learner's task involvement under different conditions of feedback. Journal of Educational Psychology, 1964, 55, 167-173.

Uhlinger, C. A., & Stephens, M. W. Relation of achievement motivation to academic achievement in students of superior ability. Journal of Educational Psychology, 1960, 51, 259-266.

Warner, D. A. & deJung, J. E. Effects of goal setting upon learning in educable retardates. American Journal of Mental Deficiency 1971, 75, 681-684.

Wechsler, D. Manual: Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1949.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Wright, W. R. Some effects of incentive on work and fatigue. Psychological Review, 1906, 13, 23-34.

Zontine, P. L., Richards, H. D., & Strang, H. R. Effect of contingent reinforcement on Peabody Picture Vocabulary Test Performance. Psychological Reports (In Press).

Appendix 1

First Day of School Information

NAME_____

DATE OF BIRTH_____

MALE OR FEMALE_____

DESCRIBE AS BEST YOU CAN YOUR MOTHER'S JOB_____

_____

_____

DESCRIBE AS BEST YOU CAN YOUR FATHER'S JOB_____

_____

_____

Appendix 2

NAME_____

PERIOD_____

1. Do you worry when the teacher says that she is going to ask you questions to find out how much you know? yes no

In the following questions the word "test" is used. What I mean by "test" is any time the teacher asks you to do something to find out how much you know or how much you have learned. It could be by your writing on paper, or by your speaking aloud, or by your writing on the blackboard.

2. Do you worry a lot <u>before</u> you take a test? Yes No

3. Do you worry a lot <u>while</u> you are taking a test? Yes No

4. Do you worry a lot after you have taken a test about how well you did? Yes No

5. When the teacher says that she is going to give the class a test, do you become afraid that you will do poorly? Yes No

6. When the teacher says that she is going to give the class a test, do you get a nervous or funny feeling? Yes No

Name_____

1. Do you mind taking school tests?

no		I don't care		yes
1	2	3	4	5

2. Did you mind taking this test?

yes		I didn't care		no
1	2	3	4	5

3. There are forty (40) questions on this test. Please write how many you think you got right. _____.

4. There were five questions on each page. Please circle the number of questions that you think you got right on each page.

1	2	3	4	5

5. When you took the test today, what did you aim for? Circle the best answer.

   a. I aimed to get a certain number right.

   b. I tried my best.

   c. I didn't try.

6. Yesterday you took Test 1 which was like this one. Do you think Test 1 was easier or harder or about the same as the test you took today?

Test 1 easier		Both tests the same		Test 1 harder
1	2	3	4	5

7. Did using the little answer card help you or hurt you on the test today?

hurt		no difference		helped
1	2	3	4	5

8. Did you feel more nervous during today's test or during the test yesterday?

more nervous yesterday		no difference		more nervous today
1	2	3	4	5