The purpose of this review was to analyze the literature on feedback in order to specify feedback variables whose further study could make the greatest potential contribution to the design of effective instructional materials and of group instructional materials. Topics considered include modes of feedback, amount of information in the feedback stimulus, frequency of feedback, immediacy of feedback, and feedback and incentives. (Author/RH)
FEEDBACK IN INSTRUCTION
A Review and Suggestions for Further Research
Norman C. Higgins


Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-71-2128. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.

Technical Report #20201
Instructional Resources Laboratory
Arizona State University
Tempe, Arizona
February, 1972
Acknowledgments

This report is one in a series of reports reviewing instructional variables under a project titled "Cues, Feedback, and Transfer in Undergraduate Pilot Training." This project was made possible by a grant from the Air Force Office of Scientific Research.

The author wishes to express his gratitude to those who contributed to the preparation of the review. Vernon S. Gerlach, Project Director, gave generously of his time in criticizing and discussing the ideas presented in the paper. Howard J. Sullivan was most helpful in suggesting ideas for the paper and in his comprehensive editing of the final manuscript.

Other members of the project staff who contributed to the production of this review include Kenneth Roberts, research associate, and Cecelia Calhoun, project secretary.

Norman C. Higgins
Arizona State University
February 1972
Introduction

The idea that learning is facilitated when the subject receives some form of feedback to indicate the adequacy of a response is a key principle in several prominent theories of learning (Hull, 1952; Skinner, 1953; Spence, 1956; Gagne, 1965). The results of numerous studies with humans and animals were responsible for the initial derivation of this principle and subsequent studies have consistently supported it. Programmed instruction, with its emphasis on learner response and concomitant potential for frequent feedback, has increased awareness and acceptance of the importance of feedback in instructional situations.

The most thorough review of the literature on feedback is Annett's (1969) analysis and discussion of over 70 years of feedback research and theory. Less comprehensive reviews limited to instructional applications of feedback have been prepared by Anderson, 1967; Briggs and Hamilton, 1964; and Gagne and Rowher, 1969. Although these reviews document the general effectiveness of feedback in facilitating learning, they also emphasize the need for further research evidence that can be directly applied to the design of instructional materials and procedures. The available research on feedback fails to provide answers to many questions regarding the most effective combinations of feedback variables in instruction.
The purpose of the present review was to analyze the literature on feedback in order to identify specific feedback variables whose further study could make the greatest potential contribution to the design of effective instructional materials and of group instructional presentations. The literature sources used in the review include Psychological Abstracts, Research in Education, Education Index, ERIC, and a bibliography prepared by the Defense Documentation Center (DDC) on the topics of feedback, large group instruction, pilot training and responder systems. The comparative dearth of good research on feedback in group instructional settings is indicated by the fact that less than a dozen of 300 reports from the DDC dealt with training in group situations. The other literature sources yielded a similarly small number of well-designed research studies on feedback in group instruction.

**Modes of Feedback**

Feedback has been studied most extensively in motor learning tasks. In such tasks, the term "feedback" has been used to refer to the information that a subject receives about the adequacy of a response as a consequence of that response. One or more stimuli under the subject's control are affected by his response, and his perception of the effects of his response constitutes feedback.

The effects of feedback in motor skills tasks have normally been studied under conditions where the only form of feedback available to the subject is the condition of the stimulus that occurs as a direct consequence of his response. This type of feedback is called
intrinsic, or task-intrinsic, feedback. Intrinsic feedback to movement of the stick in flying an airplane, for example, could be provided by the resulting changes in the turn and bank indicator, the attitude indicator, and the angle of the horizon. In several studies, the effects of adding an external visual or auditory feedback stimulus, such as a buzzer or flashing light to indicate the adequacy of the subject's response, have been investigated (Bilodeau, 1952; Morin and Gagne, 1951; Reynolds and Adams, 1953). Feedback employing an external feedback stimulus in conjunction with the task-intrinsic stimulus is called augmented feedback.

Successful performance of most tasks, whether or not they require complex motor performance, requires the acquisition of information about procedures for performing the task. This review deals primarily with feedback as it relates to the acquisition of such information. Whereas the feedback stimulus indicating the adequacy of the subject's response varies as a direct consequence of the response in performance of motor tasks, task-intrinsic feedback indicating the adequacy of the learner's response does not occur as a direct consequence of performance on cognitive learning tasks. The most common and cost-effective method of indicating the adequacy of a response in most cognitive learning tasks is to present a standard feedback stimulus irrespective of the subject's response on a given task. The feedback stimulus indicates the correct response and may also state why that response is correct. Because this review is concerned with both instruction and cognitive
learning tasks, the term "feedback" will be used to refer to any stimulus indicating the adequacy of a preceding response, and not just to task-intrinsic feedback stimuli.

Sources of Feedback Research

The research most often cited in describing the role of feedback in instruction has come from three sources of experimentation: laboratory studies of human learning, studies employing existing conventional instructional materials that in their original form do not require overt learner responding, and investigations with programmed instructional materials. Most research from the first two sources has very limited applicability to the design of instructional materials because of the difference between the materials and procedures employed in these studies and those employed in systematically designed instruction. In most laboratory studies of human learning, the learner receives no instruction before being asked to respond. The investigations by Bourne, et al. (1958, 1963, 1965), of the effects of variations in feedback on concept learning are typical of this type of study. In such studies, the learner is initially required to guess the correct response. This procedure is ineffective instructionally when contrasted with procedures in which the learner receives initial instruction designed to enable him to respond correctly (Wittrock, 1963). The effects of feedback can be very powerful under the "guessing" condition employed in most laboratory studies because the feedback stimulus constitutes the only means of determining correct succeeding responses.
However, feedback effects have not been as powerful in studies involving instruction designed to establish correct learner responses as efficiently as possible.

The limited usefulness of studies of feedback in conventional materials is due primarily to the lack of overt learner responding and of specific learning objectives for the materials. The typical procedure in such studies is for investigators to modify the materials by inserting sets of questions requiring learner responses (e.g., Angell, 1949; Hirsch, 1952; Michael and Maccoby, 1961; Sturges, 1969). However, analysis of materials of this type often reveals such instructional inadequacies as content irrelevant to the criterion tasks serving as the dependent variable in the study and insufficient instruction and practice related to these tasks. Research findings resulting from the use of such materials cannot be used as a basis for the systematic design of instruction. As noted by Anderson (1969), research on feedback in modified conventional materials has, in fact, often-produced results conflicting with those obtained in studies employing materials that are systematically designed for the achievement of specific objectives.

Because of the limited applicability to instructional design of laboratory studies and of investigations involving conventional materials, the primary source of studies used as a basis for this review was investigations employing programmed instruction. An additional source for the review was studies using non-programmed instructional sequences with
specific objectives, direct instruction and learner practice on the objectives, and a criterion test assessing learner attainment of the objectives.

**Feedback Variables in Instruction**

An analysis of the feedback literature with greatest relevance for instruction suggests several variables whose further study may have potential for contributing to the design of effective instruction. These variables include (1) the amount of information contained in the feedback stimulus, (2) frequency of feedback, (3) immediacy or feedback used, (4) the interaction between feedback and incentives for acceptable performance.

**Amount of Information in the Feedback Stimulus**

The most common explanation for the effectiveness of feedback is that it supplies information to the learner about the correct response for a particular task (Annett, 1969). However, failure to adequately differentiate between the various forms of feedback and a lack of precision in specifying the amount of information in feedback have led to a confounding of variables in many studies.

Feedback may be categorized into three common forms that provide different amounts of information. The three forms are knowledge of results (KR), knowledge of the correct response (KCR), and instructional feedback. KR, the form containing the least information, indicates only whether a response is correct or incorrect. If the response is incorrect,
KR does not indicate what the correct response is. KCR differs from KR in that KCR always indicates the correct response. Instructional feedback, the form containing the most information, indicates the correct response and provides an explanation of why the response is correct. An example of instructional feedback following the question "Which photograph illustrates the correct wingtip flight formation of the T-38 as viewed from wing position?" would be the statement:

Photograph B illustrates the proper flight formation because the wing tip of the lead plane is aligned with the star on its fuselage.

This statement is an example of instructional feedback because it explains why the correct response (Photograph B) is correct.

Studies investigating the relative effectiveness of KR and KCR in instruction typically have found that either KCR is more effective than KR or that there is no difference between the two conditions. Meyer (1960) found KCR to be superior to KR in a program designed to teach the student to spell unfamiliar words. Travers, et al. (1964), reported a similar finding in teaching German equivalents of English words. Children's rate of acquisition was increased by providing the correct response term after an error, instead of simply indicating that the response was wrong. Buss, et al. (1956), Fleming (1963), and Frase (1967) also obtained results indicating that providing the correct response (KCR) is more effective than simply indicating that a response is right or wrong (KR). In contrast, several other studies (McDonald and Allen, 1962; Merrill, 1965; Moore and Smith, 1964; Anderson,
Kulhavy and Andre, 1971) have reported no significant difference between the effects of KR and KCR on learner posttest performance.

Research comparing the effectiveness of KR and KCR would seem to be of less consequence for the design of instruction than research comparing KCR with instructional feedback. In order to present KR, materials must be computer-based, chemically treated, or presented in some other manner that permits analysis of the correctness of the learner's response. KCR, on the other hand, does not require analysis of the learner's response. Even though KCR provides more information than KR, it requires very little more space or learner effort and it can be built directly into programs with simpler and much less expensive formats. Like KCR, instructional feedback also has the advantage of being adaptable to a wide variety of program formats. However, provision of instructional feedback for a high percentage of learner responses may lengthen a program considerably, thereby resulting in increases in the program cost and in the probably amount of time required for completion of the program. Whereas the inexpensiveness of KCR and its adaptability to various program formats makes it a logical choice over KR in most instructional settings, the decision between KCR and instructional feedback, is not nearly so easy to make on an à priori basis.

Unfortunately, less research has been conducted on the comparative effectiveness of KCR and instructional feedback than on the relative efficacy of KCR and KR. The limited research on the former issue (Merrill and Stolurow, 1966) suggests that instructional feedback is more effective than KCR.
Several reviewers (e.g., Anderson, 1967; Annett, 1969) have noted that one factor influencing the effectiveness of feedback is the "nature of the learner." Specifically, it appears that feedback is more effective when the learner possesses a low level of competence with regard to the instructional task. A study by Melching (1966) indicates that, when students are given the option of requesting feedback for each response, low-ability students request it much more frequently than high-ability students. Students in Melching's study responded incorrectly on 28 percent of the frames for which they requested feedback and on only 4 percent of the frames for which they did not request it. This suggests that learners want feedback or feel a need for it when particular frames or items are difficult for them, but they do not feel the need for it when the items are relatively easy.

It seems probable that more complete feedback is more effective for difficult practice items for a given learning task than it is for easier items. If this is indeed the case, a seemingly effective procedure would be to provide feedback containing a high level of information (i.e., instructional feedback) on the initial practice items for a task and less informative feedback (KCR) on later items. That is, feedback containing more complete information could be provided when the response is being acquired, but a briefer form of feedback should suffice when additional practice is being provided in an effort to ensure maintenance.

The present state of research knowledge about information in the feedback stimulus suggests that the question "What is the best
combination of feedback to employ in a sequence of instruction?" is likely to generate more productive research than the more frequently studied question "Which form of feedback is most effective?". It appears from the research literature that potential contributions to the design of effective instruction can be generated by research efforts contrasting the effects of (1) various combinations of instructional feedback and KCR and (2) combined instructional feedback and KCR with the effects of each form used exclusively.

Frequency of Feedback

Several studies have investigated the effects of presenting either no feedback or intermittent KCR in a sequence of programmed instruction. In a sense, these studies represent an extension of the studies reviewed in the preceding section, because frequency of feedback can be considered to be a special instance or level of the independent variable "amount of information in the feedback stimulus." That is, failure to present feedback after a response creates a situation in which no information is presented to the learner. This situation, of course, constitutes an even lower level of feedback information than KR.

Both Krumboltz and Weisman (1962) and Rosenstock, Moore and Smith (1965) have found that more frequent feedback yields better learner performance on the instructional program itself but no reliable difference in posttest performance. The Krumboltz and Weisman study was conducted with college students and employed a programmed textbook on tests and measurements. The study included four feedback variations:
continuous KCR, fixed ratio KCR, variable ratio KCR, and no feedback. Rosenstock, Moore and Smith employed sixth-grade subjects using a programmed text in set theory that was modified to produce four variations in feedback: 100% KCR, 20% fixed-ratio KCR, 20% variable-ratio KCR, and no feedback. Error rate on the program was significantly lower for the group receiving 100% KCR than for the other groups. However, the authors indicate that the lower error rate in the 100% KCR group may be attributable to looking ahead and copying by some students in the group. (This same explanation, of course, cannot be completely disregarded for the Krumboltz and Weisman findings.) Nonetheless, if copying did occur, it apparently did not have major effects on other important criterion measures, since no significant differences occurred in posttest achievement, retention, or time to complete the program.

The differential effects of varying the levels and frequency of feedback in the instructional program itself have also been reported by Lublin (1965). Lublin's study involved college students using the Holland and Skinner programmed textbook, Analysis of Behavior, under differing schedules and percentages of KCR. Learners receiving no feedback took longer to complete the program and attained higher posttest scores than the group receiving 100% KCR. Lublin interpreted her results to indicate that students who did not receive KCR studied the instructional material more carefully. A similar interpretation has been offered by Sullivan, Baker and Schutz (1967) for superior performance in an instructional program and for longer program completion time by Air Force ROTC cadets who received no feedback than by cadets who received immediate KCR for each response.
Studies such as those reviewed in this section in which the presentation of the feedback stimulus is under the control of the learner tend to obscure the effects of feedback per se on learner achievement because variations in availability of feedback may alter study behaviors which also affect performance. The student receiving continuous KCR may work more rapidly and carelessly through instructional portions of a program because he knows correct answers will be presented to him after he responds. If differential achievement occurs either in the program or on the posttest, it is difficult to determine whether the achievement differences are due to the feedback variations per se, differing study behaviors, or both.

In an attempt to eliminate potential problems associated with student control over presentation of feedback, Anderson, et al. (1971), conducted a study employing computer-based instruction to require a learner response to each frame before presentation of KCR. College students were administered a programmed lesson on diagnosis of myocardial infarction under four feedback conditions: 100% KCR for each correct response, 100% KCR for each incorrect response, 10% KCR presented on a random ratio for correct responses only, and no feedback. While performance in the program itself often favors learners receiving a high amount of KCR in programs where feedback delivery is under greater learner control, it did not differ significantly for the treatment groups under the conditions in this study. Achievement on the posttest, however, favored the groups receiving KCR for correct
responses, with the group receiving 100% KCR for correct responses scoring significantly higher than the group that received no feedback. Thus, results under these conditions of greater experimental control over delivery of the feedback stimulus differed rather sharply from those obtained in studies in which presentation of feedback was under greater learner control.

As noted above, the results of many studies of the effects of frequency of feedback are difficult to interpret because variations in availability of feedback may often affect learner study behavior in the instructional program in addition to (or rather than) en route and posttest achievement. The research data on information in the feedback stimulus indicates that the "no information" condition which exists when feedback is withheld is relatively ineffective. Although the study by Anderson, et al. (1971), supports this conclusion, it is important to note that the conditions existing in the other studies reviewed on frequency of feedback may be more typical of most learning situations than were those in the Anderson study. Thus, there is the possibility that the potentially positive effects of consistent feedback may be negated by the unintended effects that it has on the study behavior of students.

Research designed to identify procedures for maximizing the potentially positive contributions of feedback and for overcoming its possible negative effects is discussed in the later section on feedback and incentives.
Immediacy of Feedback

Investigations of the effects of delaying presentation of feedback in instructional programs have yielded a variety of results. Variations in immediacy of feedback have been found to influence study behavior, achievement on in-program tasks, posttest performance and retention. Difficulties arise in determining the effects of immediacy per se on achievement because of concomitant effects of variations in immediacy on study behaviors, which influence achievement. Thus, the effects of variations in immediacy and in study behavior become confounded and present interpretation problems similar to those associated with frequency of feedback.

Schutz, Baker and Gerlach (1964) and Sullivan, Baker and Schutz (1967) have reported differences in study behavior under conditions where immediate feedback was available to learners. Chemically treated answer sheets which required the student to respond with a special pen before receiving feedback were employed with textual instructional materials in both studies. When the subject marked a correct response blank for an item, the blank turned red; when he marked an incorrect blank, it turned yellow. Subjects who initially responded incorrectly to an item under this condition continued to mark their next most preferred response choice until they responded correctly, thereby receiving KCR for each item. This procedure was designed both to provide immediate feedback and to prevent learners from looking ahead to find the correct answer before responding. However, a sizable number of the
intermediate-grade subjects in the series of investigations by Schutz, Baker and Gerlach responded initially by making only a tiny dot in an answer blank, then filling in the blank only if the dot turned the correct color. Subsequent modification of the chemical compound used on the answer blanks caused the color to emerge and spread when marked even lightly, thereby eliminating use of the "dotting" tactic. Nonetheless, differences still were observed on measures of in-program performance when the same immediate-feedback procedure was employed by Sullivan, et al. (1967), with AFROTC cadets on periodic practice tests inserted into carefully sequenced textual material. Cadets who did not receive feedback to responses on the practice tests obtained significantly higher scores on the en route practice tests than did subjects receiving immediate feedback. The experimenters interpreted these data as supporting their observations made during the instructional sessions that subjects receiving immediate feedback neglected the textual material to some degree and relied in part upon the instructional value of the immediate KCR for their responses. That the subjects receiving immediate feedback achieved some success from their procedure is suggested by the fact that their posttest achievement was comparable to that of their counterparts who did not receive feedback, even though their performance on similar and identical tasks during the program was significantly inferior.

Variations in immediacy of feedback in instructional materials have not commonly resulted in differences in posttest achievement
(Feldhusen and Birt, 1962; Brackbill, Wagner and Wilson, 1964; and Schutz, Baker and Gerlach, 1964). However, a number of investigators employing paired-associate learning tasks have found that delay of feedback facilitates retention (Brackbill, 1964; Sassenrath and Yonge, 1968, 1969; Sturges, Serafin and Donaldson, 1968). The typical finding in these studies was that there were no achievement differences on posttests administered immediately following instruction, but significant differences favoring subjects receiving delayed feedback occurred on retention measures administered five to seven days after instruction. Retention differences in favor of students receiving delayed feedback have been observed when feedback is delayed by periods ranging from five seconds (Brackbill and Kappy, 1962; Brackbill, Bravos and Starr, 1962) to as long as 24 hours (Sturges, 1969, 1970; Sassenrath and Yonge, 1968).

Since instructors and designers of instructional materials normally seek to produce relatively permanent changes in learner behavior, study of the effects of instructional variables on retention represents an area of considerable relevance for instructional designers. While the research noted above indicates that delayed feedback may indeed facilitate retention, only a few studies (Sturges, 1969, 1970) related to this topic have employed types of meaningful learning tasks and content similar to those commonly found in training and group instruction. These studies suggest that the positive effects of delayed feedback also occur with meaningful material. Useful future research in this area would include (1) collection of additional data on the effectiveness of delayed feedback on retention of meaningful instructional
material, (2) investigation of the effects of the administratively practical task of withholding feedback for responses to a series of tasks and presenting it at one later time for the entire series, and (3) study of the length of time over which the retention effect persists, if it does indeed consistently occur on instructional tasks.

In investigations involving variations in the frequency and immediacy of feedback, automated instructional procedures should be effective in reducing the confounding of effects of the feedback variable with effects of differences in learners' study behavior occurring as a result of the feedback variations. Such procedures can be used to control presentation of the feedback stimulus and thereby to prevent the learner from using tactics for discovering correct answers through feedback rather than in the instructional portion of the material. Provision of adequate incentives for achievement under both experimental and natural conditions may also reduce the differences in study behavior associated with the feedback variations and maximize their potential effects.

Feedback and Incentives

One of the most critical problems in the delivery of instruction is the establishment of desired control over learner behavior. Neither conventional printed instructional materials nor programmed materials are normally considered very stimulating by learners. A further handicap is present in many studies involving experimental instructional materials and procedures because no grade or other potential reinforcer for the learner is contingent on his performance. Thus, there is
often less incentive for the learner to perform well in experimental studies of instruction than in regular classroom instruction. Under such conditions, students often exhibit behavior that is not conducive to learning no matter how high the quality of the instructional material. Failure of learners to attend carefully to the instructional material and to the questions in instructional programs has been reported by several investigators (Anderson, Kulhavy and Andre, 1971; Dick, 1963; Schutz, Baker and Gerlach, 1964; Kress and Gropper, 1964). It is not realistic to expect variations in feedback to produce major differences in learner achievement under conditions of this type.

The facilitating effect of reinforcement on both performance and learning is well known. It has been repeatedly demonstrated that learner responding in instructional situations can be brought under the control of an effective reinforcer and that learners are highly sensitive to experimentally induced changes in reinforcement schedules and conditions (e.g., Hewett, 1967; Maltzman, Holz and Kunze, 1965; Schutz, Sullivan and Baker, 1968; Staats, Staats, Schutz and Wolf, 1962). Establishment of a reinforcement contingency for acceptable learner performance in an instructional program and/or on the posttest should produce greater control of desired learner behaviors and maximize the effects of variations in the instructional materials. For example, if an effective reinforcer were available for learner achievement, learners could be expected to attend more closely to the instructional material and the feedback and to attempt to learn as much as possible
from both sources. Learner performance under these conditions should be a much better indicator of the potential effects of variations in feedback than performance under typical experimental conditions in which there is little incentive for achievement.

Few studies have been conducted of the combined effects of reinforcement and variations in feedback on learner performance. However, two such studies conducted in 1966 with AFROTC cadets by Sullivan, Baker and Schutz (1967) and Sullivan, Schutz and Baker (1971) suggest that variations in the reinforcement contingency can produce differential feedback effects. In the first study, cadets under conditions of either immediate feedback or no feedback could earn up to $4.00 each for successful posttest performance after four one-hour instructional sessions, while a second group of cadets under the same feedback conditions received $2.50 each irrespective of performance. No differences in posttest achievement resulted from either the feedback or the reinforcement variations, but subjects who received immediate feedback exhibited less desirable study behavior as evidenced by their apparent neglect of portions of the instructional material and their significantly lower scores on en route tests embedded in the program. In the second experiment, a reinforcement contingency suggested by cadets was employed in an attempt to establish greater control over study behavior. This contingency enabled cadets to earn release from a maximum of three one-hour 7:30 a.m. close-order drill sessions for successful performances on en route tests and the criterion
test. Overall posttest achievement was significantly higher for cadets under this drill-period contingency than for the comparable sample under the monetary contingency in the earlier study. That a more powerful reinforcer may influence the effects of feedback is indicated by the fact that no differences associated with feedback occurred in either program-completion time or en route achievement under the drill-period contingency, whereas under the monetary contingency subjects who received immediate feedback completed the program in significantly less time and attained significantly lower scores on en route tests.

The potential effectiveness of incentives both for facilitating learning and maximizing the effects of feedback and other instructional variables deserves further study. The precise effects that would result from combining a relatively powerful reinforcer, such as release from scheduled but low-priority and unappealing activities, with variations in feedback are not clear. It seems possible that a reinforcer of this type might produce either of these contrasting results: (1) motivate study behavior intense enough to produce such high achievement from, say, materials containing little or no feedback that there would be no difference between achievement with these materials and with materials employing continuous feedback; or (2) maximize the effects of the continuous feedback and increase the achievement differences under the two conditions. The effects of various feedback variables under conditions of high and low reinforcement for achievement should contribute to identification of the most effective feedback procedures for use in instruction.
References


Fleming, M. L. Influence of three teaching machine factors...feedback to programmer, participation by learner, and feedback to learner...on the production and utilization of science films. Title VII Project No. 800. NDEA, U.S. Office of Education, 1963.


Merrill, M. D. Correction and review on successive parts in learning a hierarchical task. Journal of Educational Psychology, 1965, 56, 225-234.


Morin, R. E., & Gagne, R. M. Pedestal sight manipulation test performance as influenced by variation in type and amount of psychological feedback. USAF HRRC Research Note P and NS No. 51-7 (U.S. Air Force, San Antonio, Texas), 1951.


