no action

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION
SPONSORED PROJECT INITIATION

Date: 1/19/78

Project Title: The Human Self-Assessment Process

Project No: G-42-626

Project Director: Dr. Darwin P. Hunt

Sponsor: New Mexico State University; Las Cruces, New Mexico 88001

Agreement Period: From 9/21/77 Until 7/31/78

Type Agreement: Verbal (Subcontract under U.S. Army Research Institute grant)

Amount: $5,527

Reports Required: As appropriate - no formal agreement.

Sponsor Contact Person (s):

Technical Matters

Contractual Matters (thru OCA)

Ms. Betty L. Stevenson
Director, Grants & Contracts
New Mexico State University
P.O. Box 3699
Las Cruces, New Mexico 88003
Phone: (505) 646-2611, ext. 2306

Defense Priority Rating: Unknown

Assigned to: School of Psychology (School/Laboratory)

COPIES TO:

- Project Director
- Division Chief (EES)
- School/Laboratory Director
- Dean/Director—EES
- Accounting Office
- Procurement Office
- Security Coordinator (OCA)
- Reports Coordinator (OCA)
- Library, Technical Reports Section
- EES Information Office
- EES Reports & Procedures
- Project File (OCA)
- Project Code (GTRI)
- Other ______

CA-3 (3/76)

To Action Rec [illegible]

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION

Date: 1/8/79

Project Title: Human Self-Assessment and Learning

Project No: G-42-626

Project Director: Dr. Darwin P. Hunt

Sponsor: New Mexico State University; Las Cruces, New Mexico 88001

Effective Termination Date: 7/31/78

Clearance of Accounting Charges: n/a - all have cleared

Grant/Contract Closeout Actions Remaining:

TERMINATED

X Final Invoice ~~xxxxxxxxxxxxxxxx~~

_ Final Fiscal Report

_ Final Report of Inventions

~~X Govt. Property Inventory & Related Certificate~~ [illegible]

_ Classified Material Certificate

_ Other ______________________

Assigned to: Psychology (School/Laboratory)

COPIES TO:

Project Director
Division Chief (EES)
School/Laboratory Director
Dean/Director—EES
Accounting Office
Procurement Office
✓Security Coordinator (OCA)
Reports Coordinator (OCA)

Library, Technical Reports Section
Office of Computing Services
Director, Physical Plant
EES Information Office
Project File (OCA)
Project Code (GTRI)
Other ______________________

CA–4 (3/76)

THE HUMAN SELF-ASSESSMENT PROCESS

Study 1. Its Accuracy, Effects on Acquisition and a Tentative Model

Interim Report

by

Darwin P. Hunt
Department of Psychology
New Mexico State University
Las Cruces, New Mexico

Technical Representative, US Army Research Institute
Dr. Milton H. Maier

This research was supported by Grant #DAHC19-76-G-0002 from the US Army Research Institute for The Behavioral and Social Sciences. The views expressed herein are the authors own and do not necessarily reflect those of the Department of the Army.

July 1976

Abstract

The ability of individuals to assess the correctness of their own responses and the effects of such self-assessments on learning were investigated. Fifty Ss from each of 3 different groups (female and male students in an introductory psychology course and students in vocational studies) learned the names of 8 different pairs of pliers. All Ss indicated their answers by pressing buttons labelled with the 8 names. Ten Ss from each group learned the names under one of 5 treatments: (M) simply learn the names, (MK) learn the names and indicate one's own assessment of the correctness of each answer by pressing a "Sure" or a "Not Sure" button, (KM) the same as treatment MK except the order in which the self-assessment and the answer responses are made are reversed, and (MX) and (XM) simply learn the names and press a single button after or before indicating the answer.

The performance of the self-assessment task in treatments MK and KM had no statistically significant enhancement or impairment effect on the number of trials required to attain 50%, 75% or 100% correct responding; and there was no reliable differences among the 3 groups of subjects. The percentage "Sure" responses corresponded very closely with the percentage correct for all 3 groups of subjects and for both treatments MK and KM; the linear correlations ranged from 0.96 to 0.99. The order in which the answer and the self-assessment responses had to be made did not significantly affect the ability to discriminate between correct and wrong responses.

A tentative model of the human self-assessment process is outlined. Based upon an inspection of the response latencies under various treatments it seems reasonable to believe that the covert selection of an answer response and of a confidence in its correctness is sequential in that order. Some implications of the proposed model of the human self-assessment process are indicated; for example, the effects of feedback on learning depends upon the confidence the learner has in the correctness of a response; and the uncertainty with which an individual anticipates

consequences of his own responses depends upon both the real-world uncertainty and the uncertainty of the individual's internal model of the real-world.

Foreword

I wish to acknowledge my great appreciation to two graduate assistants, James Davis and Gus Thomason for their unrelenting help in constructing the apparatus, selecting and preparing stimulus pictures, writing computer programs, scheduling subjects, collecting and analyzing data, arranging for payment to subjects, making suggestions on a broad variety of topics and the performance of many other tasks. I especially wish to thank Mr. Davis who, partly by reason of his office being located next to mine, but mainly because of his abilities and willingness, handled more than his fair share of the details and did so with excellence.

Also, I wish to thank Dr. Ronald Downey who served as Technical Representative for the Army Research Institute during the initial portion of this grant, for his help in the selection of the stimulus and response terms used in the learning task.

Table of Contents

List of Figures

List of Figures (Continued)

List of Tables

List of Appendixes

Introduction

An instructor, or surrogate instructor (computer), requires suitable feedback information concerning the student's momentary state of knowledge if the instructor is to perform his task most effectively and if the students are to obtain the maximum benefit from the instructional experience. A group instructional situation may be viewed as an instructor (machine)-students system as shown in Figure 1. The notion is that the instructor presents material (verbally, visually, etc.)

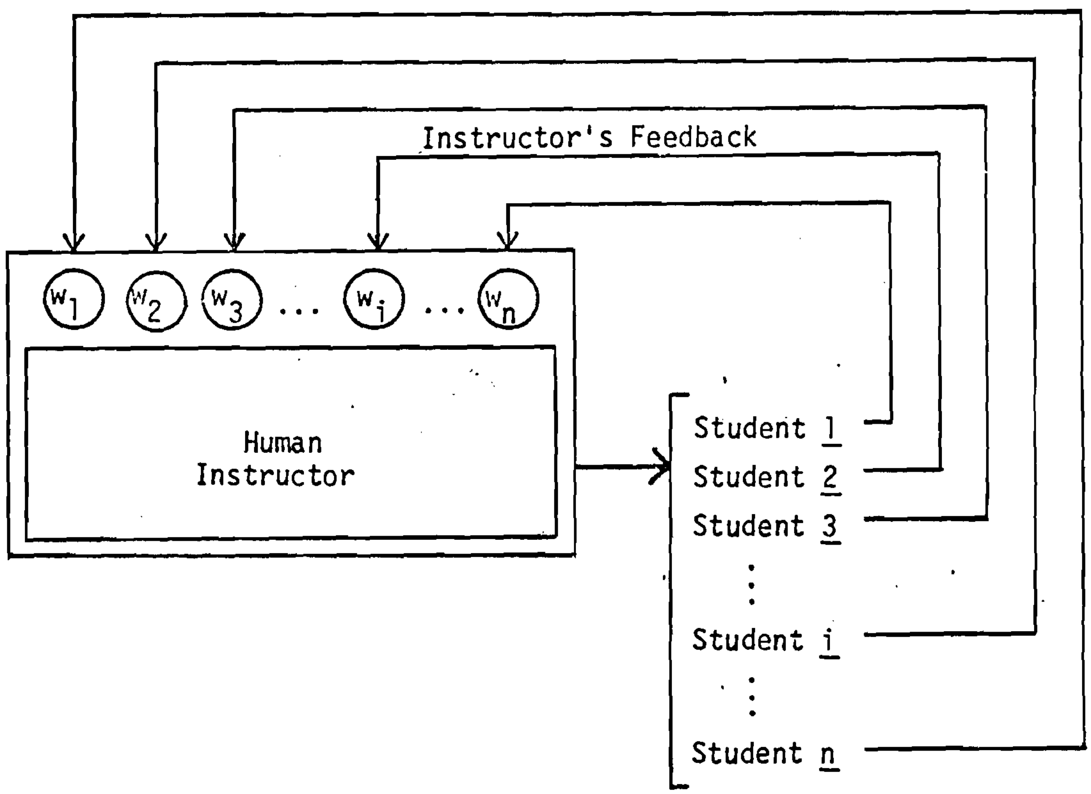


Fig. 1. A representation of an instructor (computer)-student system.

to the students, who then behave in certain ways. The students' observable behaviors serve as feedback data to the instructor which are employed by him to make inferences about the actual state of knowledge of the students relative to the state of knowledge that the instructor expects them to occupy at that time. Based upon these inferences the instructor may modify his presentation of the material to-be-learned in an attempt to reduce any discrepancy between the inferred state of knowledge and the expected state of knowledge. This reiterative process presumably continues until the student' performance is such that the instructor infers that the material, task, etc., has been learned; or the instructional process may be terminated for other reasons.

It is assumed that the feedback information provided to the instructor must be appropriate, e.g., of sufficient validity and reliability; and it should be available to the insturctor at a sufficiently high rate so as to permit the instructional presentation to be modified frequently enough. Typical testing procedures represent the traditional approach to this problem.

Another possibility for obtaining on-line feedback is to ask each learner to assess his own state of knowledge continually, or in some time-sharing fashion while he is engaged in the task of learning. The use of such self-assessment performance assumes that the learner is capable of identifying his own state of knowledge in relation to the momentary instructional goals and is capable of communicating this to the instructor in a useful fashion. As is depicted in Figure 2, this requires that the learner (a) be capable of identifying what he should "know", (b) be capable of identifying what he does "know", (c) be able to compare the two, _a_ and _b_ and (d) convey his conclusion to the instructor.

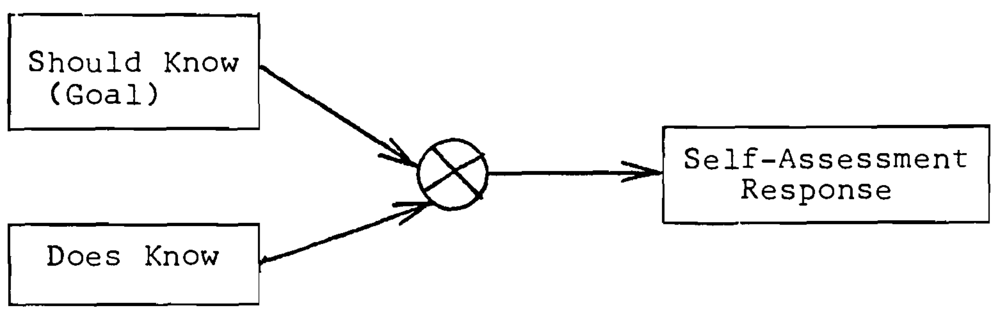


Fig. 2. Self-assessment in a dynamic learning situation.

To be of practical use it would also be desirable that performance of the secondary task of self-assessment does not unduly interfer with the primary task of learning.

The ability of individuals to assess their own knowledge state has been of interest for some time (Woodworth and Schlosberg, 1954), but is considered from a point of view that makes the work of only indirect relevance to instructional feedback systems. Some more recent work (Hart, 1966) was addressed to the "Feeling-of-knowing (FK)" experience and to determining whether this experience is an accurate indicator of memory storage. The method employed involved having a subject (S) indicate FK or FNK (a feeling-of-not-knowing) for items in an experimental list to which he had been exposed, but which the S could not recall. Then S was given a recognition test on these unrecalled items. Based upon a comparison of the percentage of correct recognitions for the FK (63%) and FNK (47%), it was concluded that the FK and FNK judgments can act as storage indicators.

Hunt (unpublished, 1965) conducted an experiment in which Ss were required to learn a fixed series of 20 randomly selected single digits (1 to 9). One group of 10 Ss simply learned the series by operating numbered toggle switches to anticipate correctly which digit would appear next in a display window. An experimental group had the additional self-assessment task of indicating, prior to each digit anticipation, whether he "knew" or "didn't know" the next digit. Three findings of this unpublished study were:

(1) the experimental (self-assessment) group learned the task in significantly fewer trials (11.8 vs 20.8 trials)

(2) Ss tended to be "optimistic," i.e., the mean percentage "know" responses exceeded the actual percentage correct; and there were large differences among Ss with the slower learners being the most "optimistic."

(3) the self-assessment responses were reasonably accurate indicators of the correctness of their answers. The conditional probabilities $P(C|K)$ and $P(C|DK)$ were calculated for each S and are shown in Table I.

Table I. Mean Number of Trials-To-Criterion For The Three Groups and The Five Experimental Treatments.

Subject	P(C|K)	P(C|DK)
AH	.933	.200
DW	.912	.033
TK	.895	.297
BK	.873	.125
JJ	.846	.233
GK	.842	.235
JH	.832	.173
SL	.779	.077
RD	.748	.190
DL	.747	.222
	.841	.178

Since only nine stimuli were employed, the P(C) by chance alone is 0.11, compared to the obtained $P(C|K) = 0.84$ and the obtained $P(C|DK) = 0.178$. The $P(C|DK)$ is not significantly different from chance performance and the $P(C|K)$ is significant ($p < .001$).

Lichtenstein and Fischoff (1976) recently reported on a series of experiments in which Ss were given test items with two alternatives. Ss were required to select one of the two alternatives and, then, indicate with a number from 0.5 to 1.0 the probability that their choice was correct. Over a large number of different groups of Ss and different circumstances, they found a correlation of 0.91 between the percentage correct and the mean probability assigned. A casual inspection of the skewed distribution of Ss probability assignments suggests that the use of the median may have produced an even higher correlation value.

The main purpose of the present study was to experimentally determine whether the performance of a self-assessment task interfers with, has no effect, or enhances the performance of a primary learning task. It was a surprise to find in the preliminary study reported above that acquisition of a serial learning task was enhanced in the group which engaged in self-assessment. However, the design of the preliminary study involved possible confounding such that the finding could be attributed to other factors such as distribution of practice. On the other hand, a recent study (Wade, 1974) found, "...that self-monitoring can significantly affect response rates. The Ss who engaged in cummulative recordings of their... responses attained both a greater number of correct responses and a greater proportion correct than did Ss who. ..did not self-record (p. 248)."

A second purpose of the present study was to evaluate the accuracy of self-assessments at different levels of acquisition. In the preliminary study employing a serial digit learning task the Ss were found to be "optimistic." However, the finding of "optimism," e.g., that the mean confidence expressed by subjects in the correctness of their response exceeds the actual percentage correctness of their responses, apparently depends upon several factors. For example, Levine et al. (1974) found that subjects performing a complex decision task tended to be "pessimistic." Their subjects were required to select one of eight possible hypotheses when they were given unreliable data from three different sources; the rate and pacing of the data presentation to the subjects and the irreversibility of the decisions were the independent variables. Each time a subject made a decision regarding which single hypothesis he believed to be true, he was required to assign a confidence (0-100%) to each of the eight hypothese (so that the sum of the confidences assigned to all eight hypotheses was equal to 100%). They report, "Overall, subjects' confidence ...was roughly 10 percentage points less than justified by empirical accuracy (p. 391)."

Based upon administering a number of two-alternative multiple choice tests to a large number of different groups, Lichtenstein and Fischhoff (1976) suggest that generally people are not very accurate; with difficult items, assessors are "optimistic" and with easy items, they are "pessimistic." They also point out that Pitz (1974) predicts that the "optimism" will decrease as knowledge increases. The present study should help to clarify some of these questions.

Another purpose of the present study was to collect data relevant to the question of whether the self-assessment process and the answer selection process are serial or parallel processes. This hypothesis can be explored by reversing the order in which the subjects are required to produce the self-assessment response and the answer response. Generally, if the processes are parallel the the order in which the two responses are required to be produced by a subject should have no effect on the response latencies. If the order in which the responses must be executed does have an effect on the response latencies, then the direction of the effect should reflect the serial order of the internal processing.

Finally, it is intended to develop a tentative model of the human self-assessment process.

Method

<u>Subjects</u>. One hundred fifty students at New Mexico State University served as subjects (<u>S</u>s); 50 male students (PsM) and 50 female students (PsF) in Introductory Psychology; and 50 students (25 male and 25 female) from the two-year vocational studies program (VOC). The original intent was to test 50 <u>S</u>s in each of one-year vocational studies, two-year vocational studies and military reserve programs (Category III), but various circumstances prevented this from being accomplished. Each subject served in one one-and-one-half hour session. The vocational students were paid $1.50; the Introductory Psychology students were given a choice of either being paid $1.50 or having their participation satisfy a portion of their course requirement. In addition, all subjects were informed that the individual in their group who learned the task the quickest would be awarded a $25.00 prize.

<u>Primary Learning Task</u>. The primary task of <u>S</u> was to learn the correct names of 8 different pairs of pliers, line drawings of which were rear projected on a screen. The line drawings were constructed based upon a review of the pictures of pliers contained in military tool catalogues. The line drawings were composed by combining two different plier heads (a short broad head and a long slender head) with two handle shapes (symetrically curved and nonsymetrically curved) and the handles were either cushioned or uncushioned. These 2 x 2 x 2: eight different pictures of pliers served as the stimuli for the paired associates learning task. The response terms of SHAPE, BEND, FORM and TWIST were initially assigned randomly to the four long slender-headed pliers and the terms SPLIT, CUT, CLIP and SNIP were assigned randomly to the four short broad-headed pliers. Once assigned these names were the same for all <u>S</u>s in all conditions throughout the experiment. The line drawings of the eight pliers and their assigned names are shown in Appendix A.

<u>Apparatus</u>. A diagram of the apparatus is shown in Figure 3. The teletype permitted instructions to be given to a PDP-8E computer and to print data and, also, punch it on paper tape. The computer controlled the presentation of (a) an easily heard tone through Telex 1200 earphones, (b) the stimulus pictures for 6 second and (c) after 1.5 second delay, a knowledge-of-results, KR, slide which contained the stimulus along with its correct name for 4 seconds. The stimulus pictures and the KR slides were rear projected on a 11.4 cm x 12.7 cm. screen at a viewing distance of approximately 60 cm.

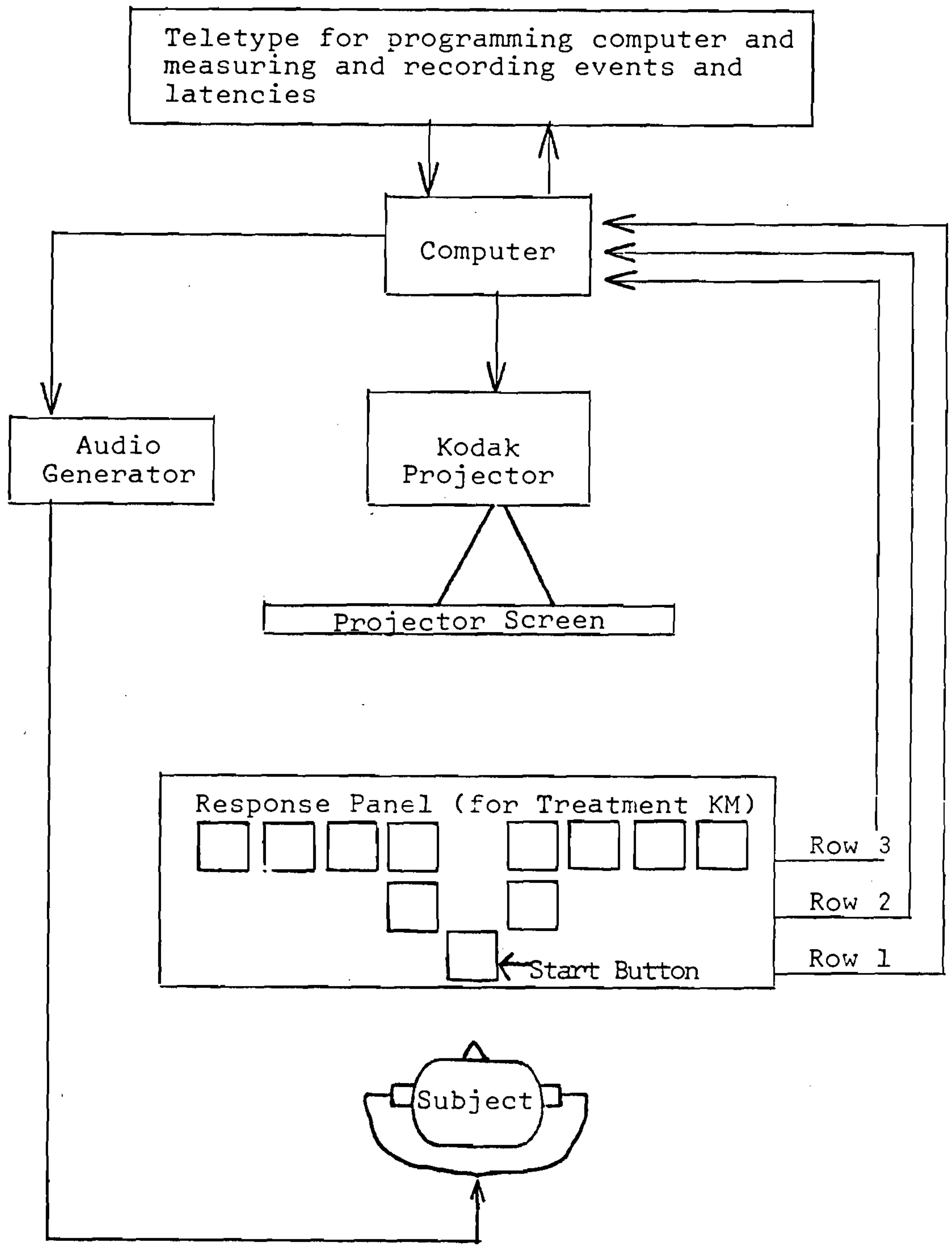


Figure 3. Block diagram of the apparatus.

A 76.2 cm. x 45.7 cm. response panel was laterally centered 30.5 cm below the bottom of the projection screen. The panel was generally horizontal but the front edge was tilted approximately 20 degrees downward to be more normal to Ss line of vision and more convenient manually. Internally lit 2.5 cm. x 1.9 cm. high response buttons on the panel could be arranged in five different ways (see Appendix B) each corresponding with one of the five different treatments involved in this experiment. For all button arrangements a START button was centered laterally and 19.3 cm. from the front edge of the response panel. The rows of buttons were separated 7.6 cm. on center. The buttons were laterally separated 1.3 cm. edge-to-edge, except the buttons next to the center line were separated by 5.1 cm.

A timer was used to measure to the nearest 0.002 second the following latencies:

Dwell Time 1 (DW1): the time from the onset of the stimulus picture to the release of the START button

Movement Time 1 (MT1): the time from the release of the START button to the activation of the next button in Row 2

Dwell Time 2 (DW2): the time from the activation of the button in Row 2 until its release

Movement Time 2 (MT2): the time from the release of the button in Row 2 to the activation of the button in Row 3

Dwell Time 3 (DW3): the time from the activation of the Row 3 button to its release

Each of these latencies as well as the identity of the buttons activated were recorded for each stimulus presentation.

Procedure. The primary task of all Ss was to learn the names of the pliers -- and they were encouraged to do so as quickly as they could both by the instructions and by informing them of the $25.00 prize. The names of the pliers were considered learned when S could go through the list twice consecutively with no error.

It was necessary for S to have the START button depressed at the time the stimulus was scheduled to be presented; otherwise the stimulus slide would not be presented and any responses would be disqualified. To alert S that a stimulus was about to be presented, a 100 millisecond tone was presented 1.5 second prior to the stimulus presentation; then if S did not have the START button depressed 1.0 second prior to the scheduled stimulus presentation the tone came on and remained on until either the START button was depressed or 1.0 second had elapsed.

The Ss in the experimental treatments MK and KM were required to assess the correctness of the M responses each time a stimulus picture was presented. This was accomplished by pressing one of two 2.5 cm. x 1.9 cm. buttons labelled SURE and NOT SURE. The button arrangements on the panel matched the order in which the M and K responses were required (see Appendix B).

Two other treatments MX and XM were included as control conditions for treatments MK and KM, respectively. The rationale for their inclusion is that any differences between treatment M and MK/KM in learning might be due to the motor responding rather than the self-assessment task.

Ten Ss were nonsystematically selected from the 50 Ss in each of the three groups and assigned to one of the five treatments:

M: Perform only the learning task, making M responses in Row 2

MK: Make the M responses in Row 2 followed by self-assessment (K) responses in Row 3

MX: Make the M responses in Row 2 followed by pressing the single button (X) in Row 3

KM: First make a K response in Row 2 followed by an M response in Row 3

XM: First press the single button X in Row 2 followed by an M response in Row 3.

Subjects were run at five specific times of day (0800, 0915, 1030, 1700 and 1815) on Mondays through Fridays by two experimenters using two different orders of answer buttons on the panel (see Appendix C). The experimental design was such that an equal number of Ss from each group-treatment combination was run at each time of day (2 Ss); and on each day of the week (2 Ss); and by each of the two experimenters (5 Ss); and using each of the two button orders (5 Ss).

After S was seated, instructions appropriate for his assigned treatment (Appendix D) were read to him which (a) informed him that his task was to learn the names of 8 different pliers, (b) pointed out the critical differences in the plier heads, handle shapes and cushioning using enlarged stimuli, (c) informed him of the various buttons on the panel, their functions, etc., for the treatment under which he was being tested, and (d) informed him of other details, e.g., S must have the START button depressed at the time scheduled for the stimulus presentation and he has 6 seconds to press an answer button and, if his treatment condition requires, press a self-assessment button.

Before the testing session began, the 8 pictures of the pliers along with their correct names were projected one time each for 5 seconds. Then the session immediately began with the projection of a single circular black dot for 3 seconds; the 100 msec. warning tone sounded, S depressed the START button and the stimuli were presented one at a time and S was required to respond as indicated by his assigned treatment. At the end of the series of 8 stimuli there was a 5½ second delay, the dot slide was presented, and the 8 stimuli were again presented one at a time, but in a different order. Seven different orders were used and these orders were cycled through until S had learned the names. If S had not learned the names of the pliers by the 40th trial then the session was ended, the data of the S was declared unacceptable, and another S was run on a subsequent week under the same conditions (day of week, time of day, experimenter, etc.) Some modest relaxation of these demands were made later, but they were such that, in the judgment of the investigator, they would have little affect on the findings of the experiment.

Results and Discussion

Effects of Self-Assessment on Acquisition.

The mean number of trials required to attain three different levels of acquisition (low, medium and high) was calculated for each group of subjects under each of the five treatments. The three acquisition levels are operationally defined:

Low: the first series (trial) on which 4 or more of the 8 answers were correct and at least 2 correct answers occurred on each subsequent trial.

Medium: the first trial on which 6 or more correct answers were made and at least 4 correct answers occurred on each subsequent trial.

High: the first of two consecutive trials on which no errors were made.

An inspection of the mean number of trial-to-criterion presented in Table II suggests that requiring Ss to assess the correctness of their responses did not enhance the acquisition of this task as was hypothesized. This observation is supported by an analysis of variance of the trials-to-criterion measures which indicates that neither the main effect of treatment nor any of the interactions involving the treatment was statistically significant. Also, there was no statistically reliable difference among the three groups (female and male introductory psychology students and vocational students) in terms of the number of trials required to learn the material under any of the five treatments.

On the other hand, this analysis of variance may be viewed as indicating that the self-assessment task did not interfer with the primary task of learning. This finding suggests that the learner can perform the simple self-assessment task of indicating "Sure" or "Not Sure" without impairing acquisition in this kind of learning task. It is expected that the acquisition may be impaired as increasingly precise self-assessment responses are requested of the learner; this will be explored in the next experiment.

Table II. Mean Number of Trials-To-Criterion for the Three Groups (Females in Introductory Psychology, Males in Introductory Psychology, and Vocational Students) and the Five Experimental Treatments. Each mean is based upon 10 subjects.

Acqusition Level	Group	Treatment				
		M	MK	KM	MX	XM
High	PSf	18.2	22.0	20.2	18.7	17.9
	PSm	16.7	15.4	20.0	21.3	16.2
	VOC	17.4	20.1	22.2	20.9	20.2
	Mean	17.4	19.2	20.8	20.3	18.1
Medium	PSf	12.8	13.2	11.0	11.8	10.1
	PSm	9.9	9.2	11.6	13.5	9.6
	VOC	11.0	14.5	12.2	13.0	12.4
	Mean	11.2	12.3	11.6	12.8	10.7
Low	PSf	7.7	8.9	7.6	6.3	6.6
	PSm	5.3	5.1	6.8	8.2	3.7
	VOC	5.7	8.1	7.7	7.4	8.8
	Mean	6.2	7.4	7.4	7.3	6.4

Accuracy of Self-Assessment Responses.

The conditional probabilities, p(Correct|Sure) and p(Correct|Not Sure), were calculated for the three groups of Ss under treatments MK and KM separately for the first third, middle third and last third of the trials. These mean p(C|S) and p(C|NS) values are shown in Table III. During the first third of the trials a few subjects did not make any "Sure" responses; and as acquisition progressed several subjects did not make any "Not Sure" responses. Thus, the mean values are based on differing number of Ss. Never-the-less some regularities may be observed. For example, the p(C|S) values increased considerably (0.5 to 0.9) from the first to the last third of the trials -- and the p(C|NS) almost doubled (0.25 to 0.5 or more).

Also, it may be noted in Table III that both the p(C|S) and p(C|NS) values are usually higher under treatment KM than under treatment MK. For p(C|S) the difference between MK and KM is small but consistent; for p(C|NS) at the medium and high acquisition levels the differences between MK and KM are fairly large.

A general notion of the accuracy of self-assessment responses might also be obtained by inspection of the percentage "Sure" responses and percentage "Correct" responses as a function of acquisition trials. In Fig. 4 (a to f) the mean percentage "Sure" and percentage "Correct" responses as a function of (Vincentized) trials are shown separately for treatments MK and KM and for each of the three groups of Ss. Since the self-assessment curve is usually at or below the percentage correct, it is clearly not possible to view these Ss in this task as being overconfident or "optimistic." In only two (male psychology students under both treatments MK and KM) or three (Vocational students under treatment MK) of the six curves is the self-assessment curve above the answer curve during even the first part (20-40%) of the acquisition. Such "optimism" early in acquisition only is at least consistent with Pitz's (1974) view that "optimism" decreases with knowledge.

Table III. The Mean p(Correct|Sure) and p(Correct|Not Sure) Values for the Treatments MK and KM at Each of the Three Acquisition Levels for Each of the Three Groups of Ss. The number of Ss upon which the mean value is based is shown in parentheses.

		Treatment MK		Treatment KM	
Acquisition Level	Group	p(C|S)	p(C|NS)	p(C|S)	p(C|NS)
High	PsF	0.87(10)	0.73(6)	0.88(10)	0.75(9)
	PsM	0.89(10)	0.45(7)	0.90(10)	0.81(7)
	VOC	0.89(10)	0.36(7)	0.91(10)	0.51(6)
	Mean	0.88	0.51	0.90	0.69
Medium	PsF	0.76(10)	0.33(9)	0.83(10)	0.44(9)
	PsM	0.79(10)	0.39(10)	0.82(10)	0.49(10)
	VOC	0.75(10)	0.43(10)	0.75(10)	0.49(9)
	Mean	0.77	0.38	0.80	0.47
Low	PsF	0.53(10)	0.22(10)	0.53(8)	0.25(10)
	PsM	0.43(10)	0.30(10)	0.50(10)	0.23(10)
	VOC	0.55(9)	0.25(10)	0.64(9)	0.33(10)
	Mean	0.50	0.26	0.56	0.27

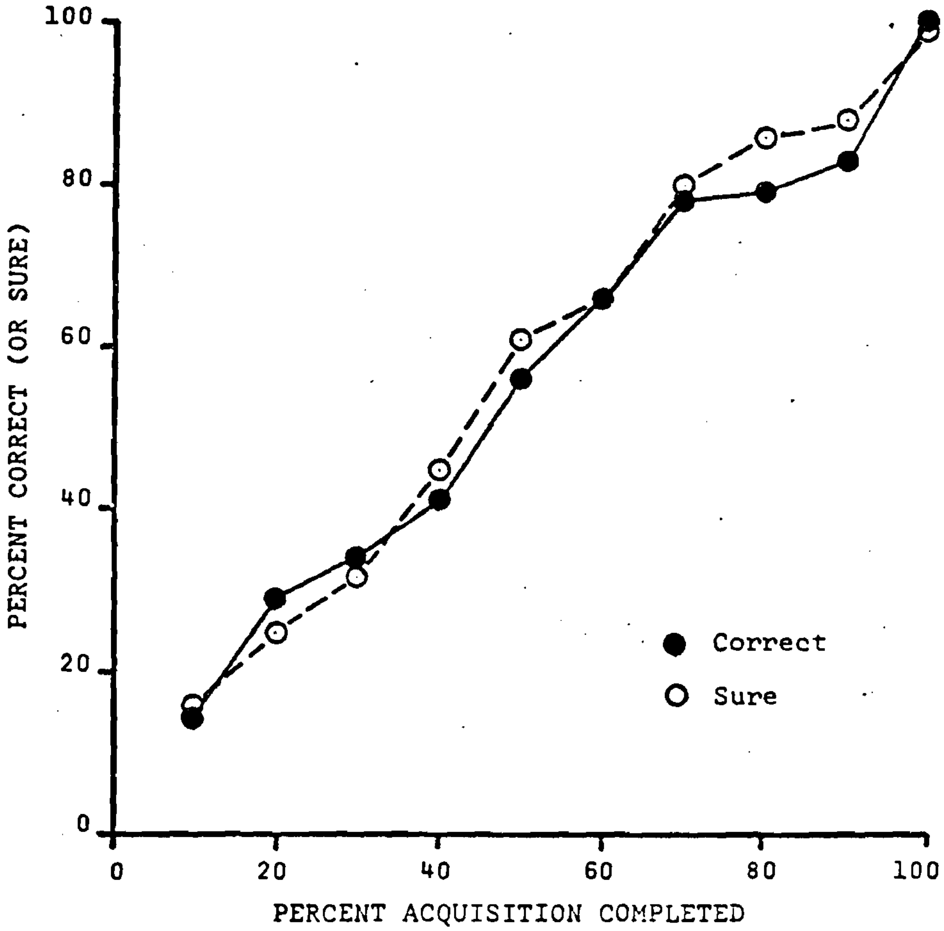


Figure 4(a). The mean percentage "Sure" responses and "Correct" responses as a function of trials (Vincentized) for females in Introductory Psychology under Treatment MK.

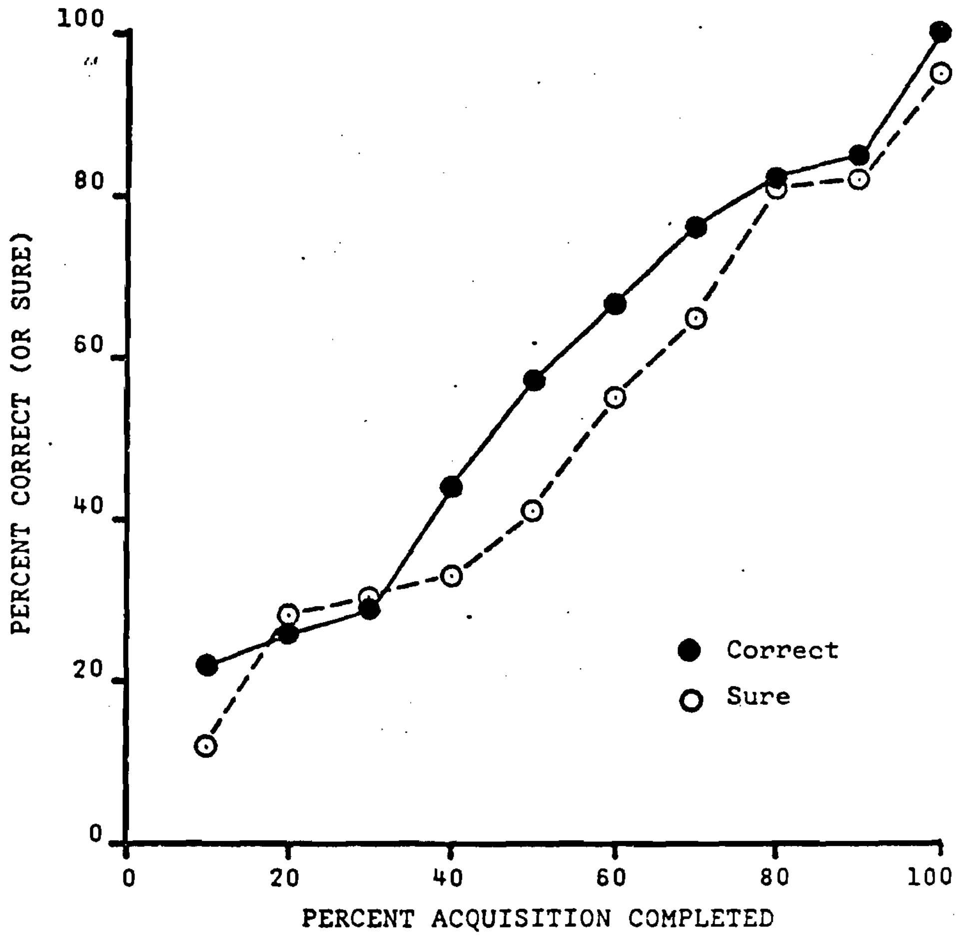


Figure 4(b). The mean percentage "Sure" responses and "Correct" responses as a function of trials (Vencentized) for females in Introductory Psychology under Treatment KM.

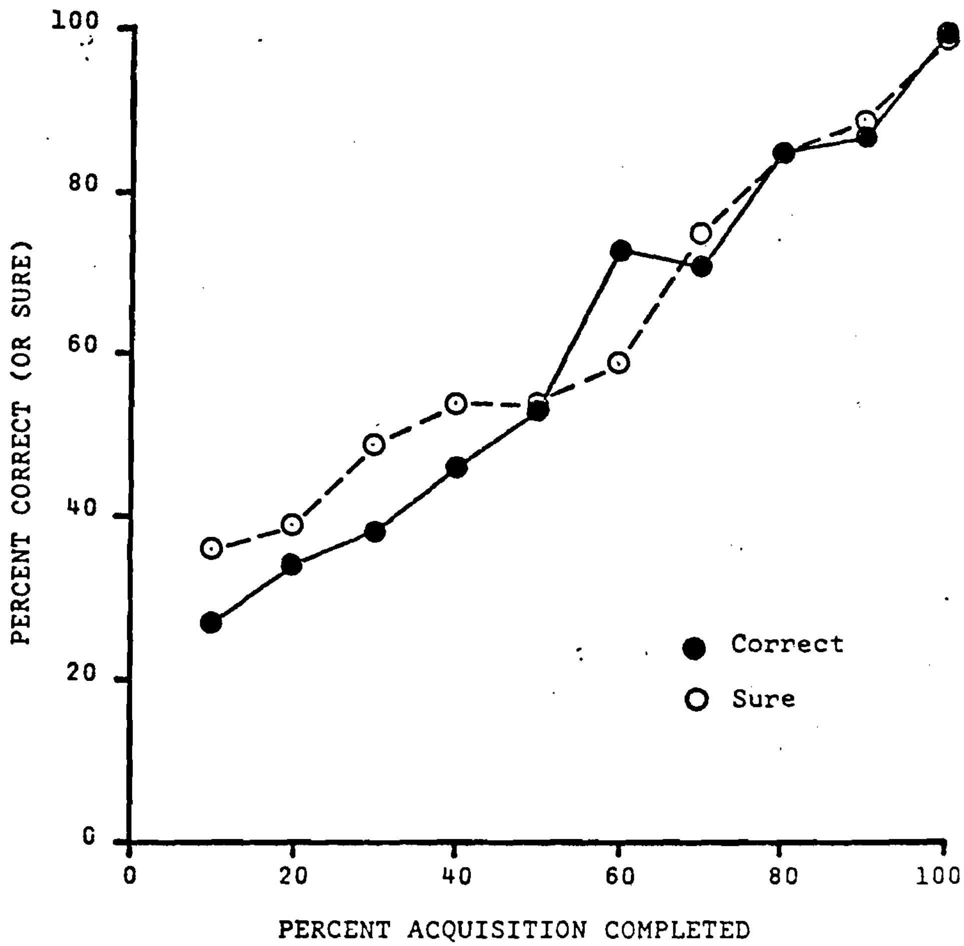


Figure 4(c). The mean percentage "Sure" responses and "Correct" responses as a function of trials (Vincentized) for males in Introductory Psychology under Treatment MK.

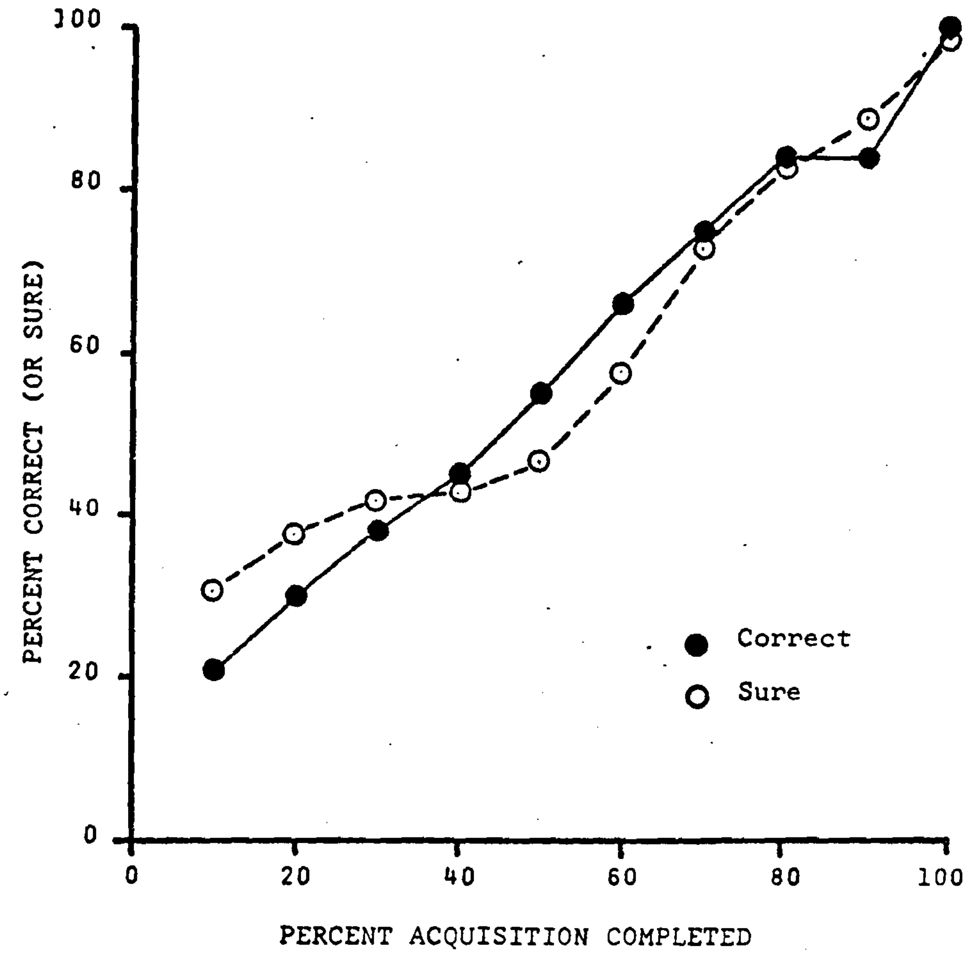


Figure 4(d). The mean percentage "Sure" responses and "Correct" responses as a function of trials (Vincentized) for males in Introductory Psychology under Treatment KM.

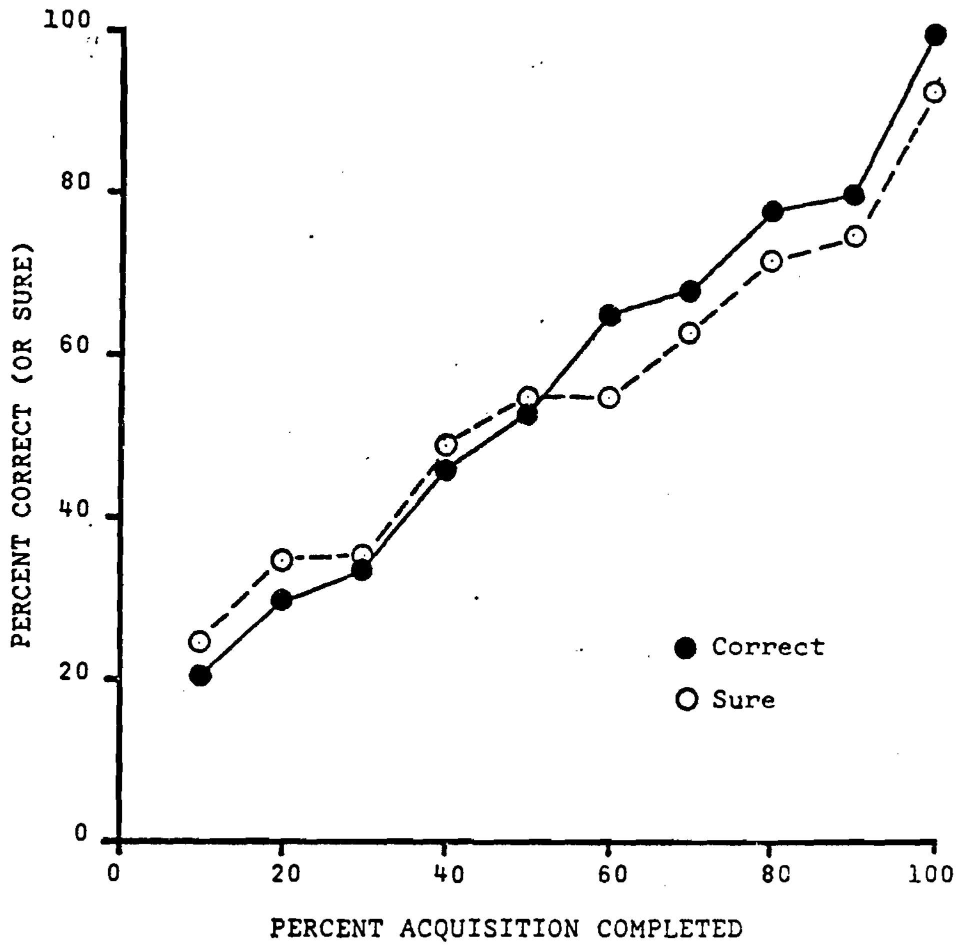


Figure 4(e). The mean percentage "Sure" responses and "Correct" responses as a function of trials (Vincentized) for Vocational students under treatment MK.

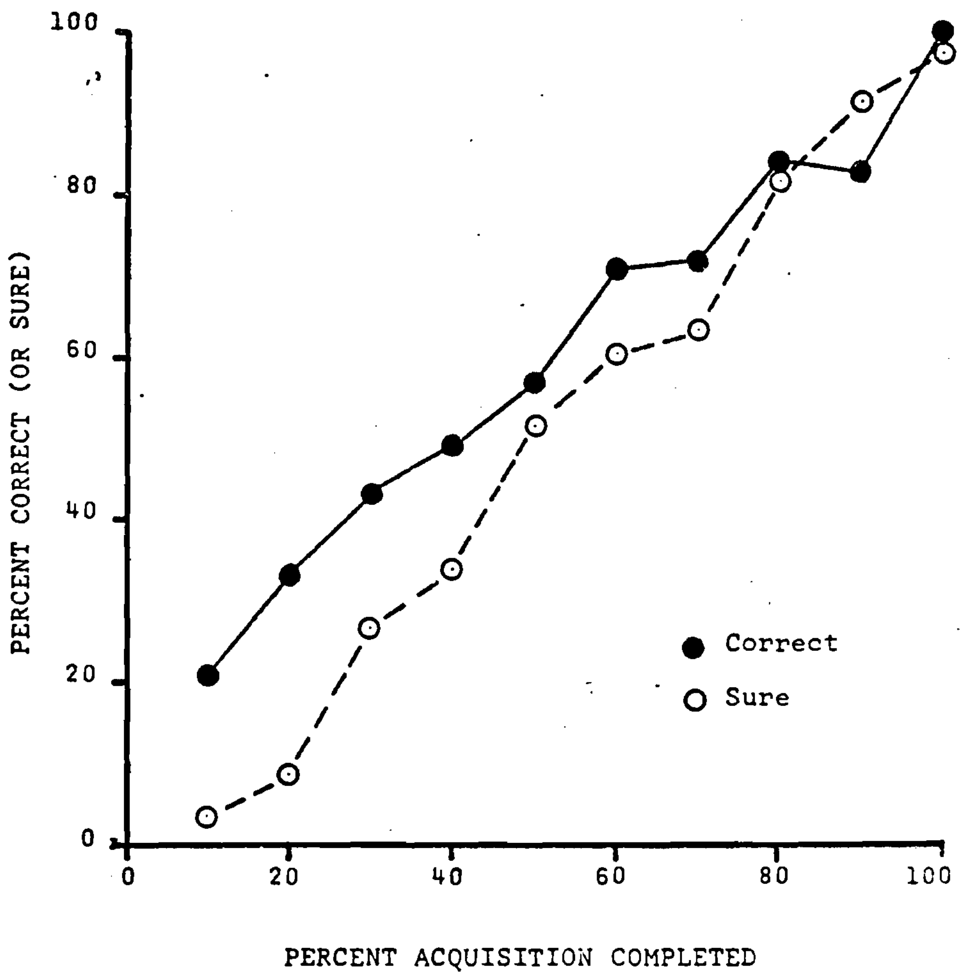


Figure 4(f). The mean percentage "Sure" responses and "Correct" responses as a function of trails (Vincentized) for Vocational students under treatment KM.

A more direct indication of self-assessment accuracy may be obtained by plotting the percentage "Sure" responses as a function of the percentage "Correct" responses. These functions are shown in Fig. 5 (a to f). The notion is that an infallible self-assessor would be represented by a line from the (0.0) point extending diagonally to the (100,100) point -- much like the "perfectly calibrated assessor" of Lichtenstein and Fischhoff (1976). In Fig. 5 the values above the diagonal represent overconfidence or "optimism" and values below the line represent underconfidence or "pessimism." The linear correlation values and the regression equation (slope and y-intercept) were calculated for each group under each treatment; these values are given along with each function in Fig. 5. The infallible self-assessor would have a correlation value of +1.0, a slope of +1.0 and a y-intercept value equal to zero. The actual correlations ranged from 0.96 to 0.99, the slopes ranged from 0.8 to 1.3 and the y-intercepts ranged from approximately -28 to +13. Some of the interpretations which are suggested by these measures are:

1. There is a high degree of linear association between the proportion of answers which are correct and the proportion of answers about which the learners said they were "sure" of the correctness.

2. The slope values of 0.8 to 1.3 in the regression equation indicate that the proportion of answers which S judges to be correct increases 8 to 13% for every 10% increase in actual proportion correct, e.g., as the material is learned.

3. The y-intercept value can be taken as an indication at least of the extent of overconfidence (a positive intercept value) or underconfidence (a negative intercept value) at low levels of correctness or knowledge such as during the initial acquisition stages.

Another approach to estimating the accuracy of self-assessment responses is to calculate a d' measure as an index of the degree to which S is able to distinguish between (a) having the correct answer stored, and retrieving it, or (b) not having the correct answer stored, or at least not retrieving it. The main concepts of this signal detection approach which are relevant to the present analysis are depicted in Fig. 6.

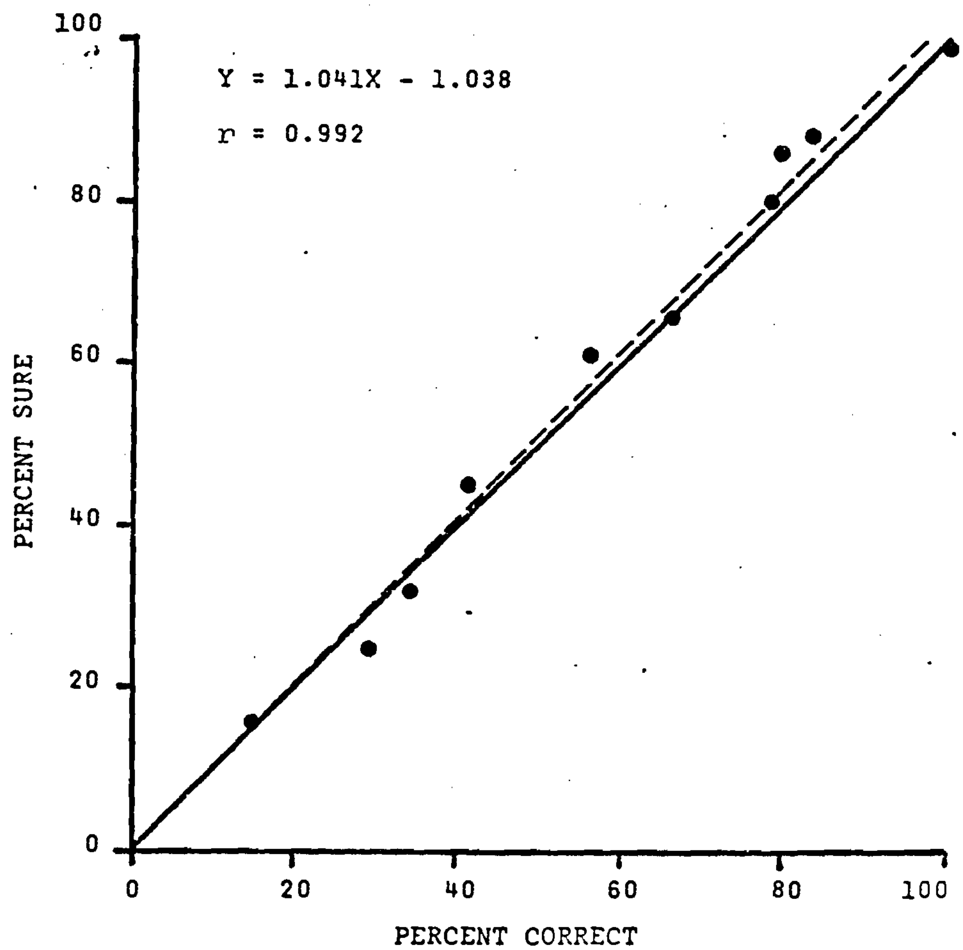


Figure 5(a). The mean percentage "Sure" responses as a function of the mean percentage "Correct" responses for females in Introductory Psychology under Treatment MK.

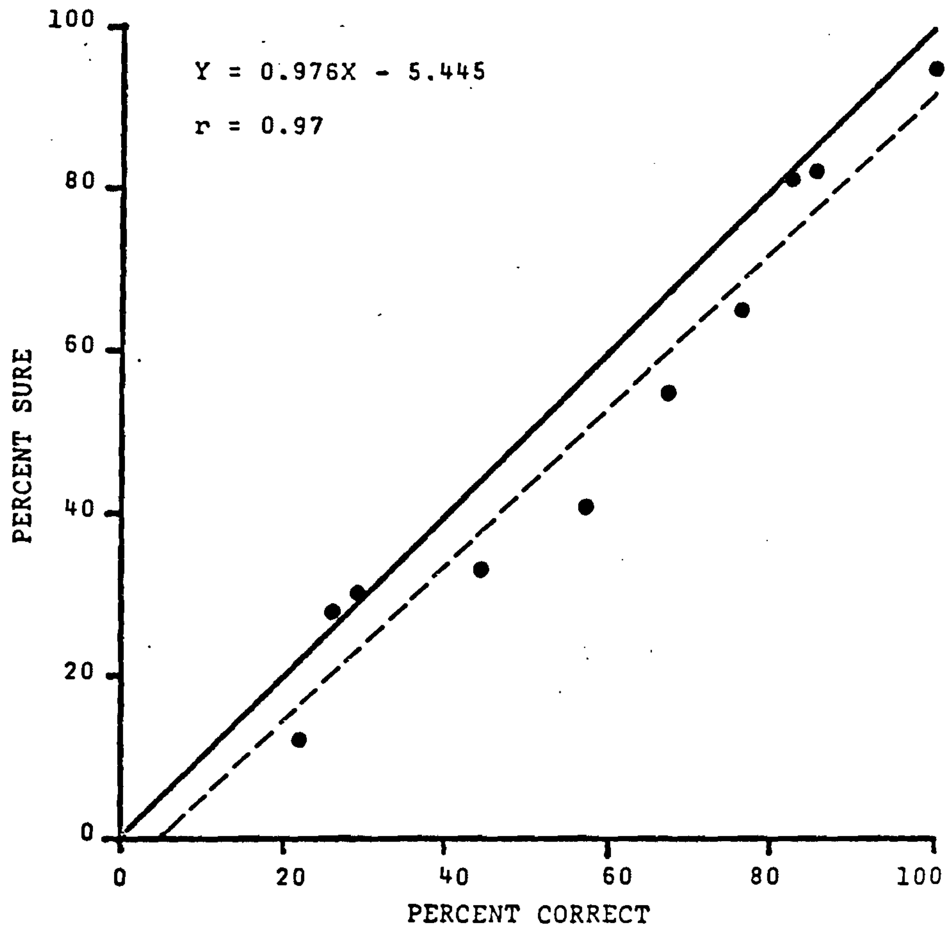


Figure 5(b). The mean percentage "Sure" responses as a function of the mean percentage "Correct" responses for females in Introductory Psychology under Treatment KM.

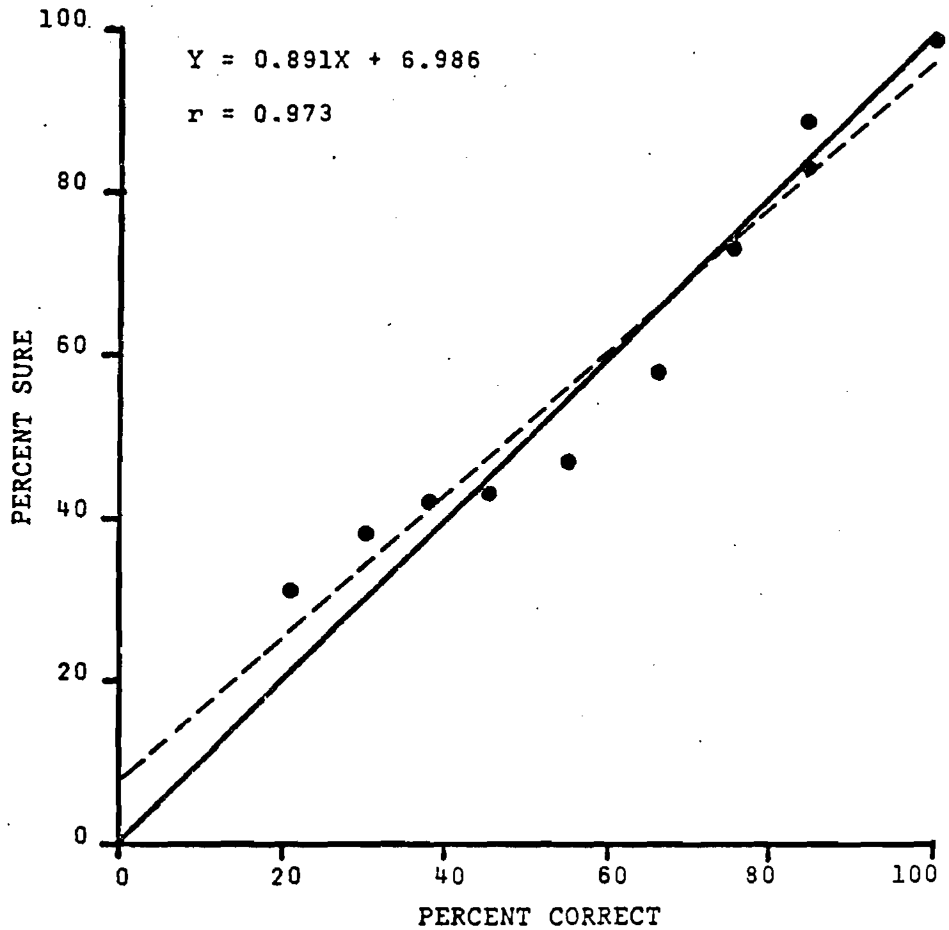


Figure 5(d). The mean percentage "Sure" responses as a function of the mean percentage "Correct" responses for males in Introductory Psychology under Treatment KM.

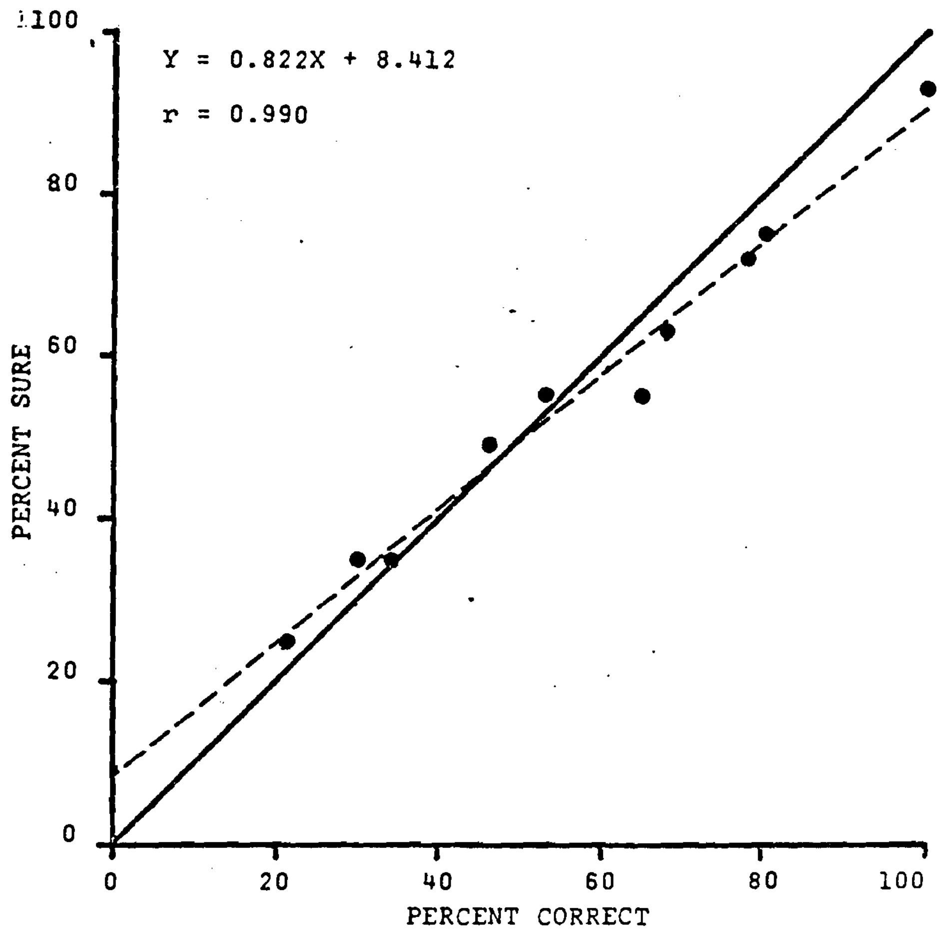


Figure 5(e). The mean percentage "Sure" responses as a function of the mean percentage "Correct" responses for Vocational students under Treatment MK.

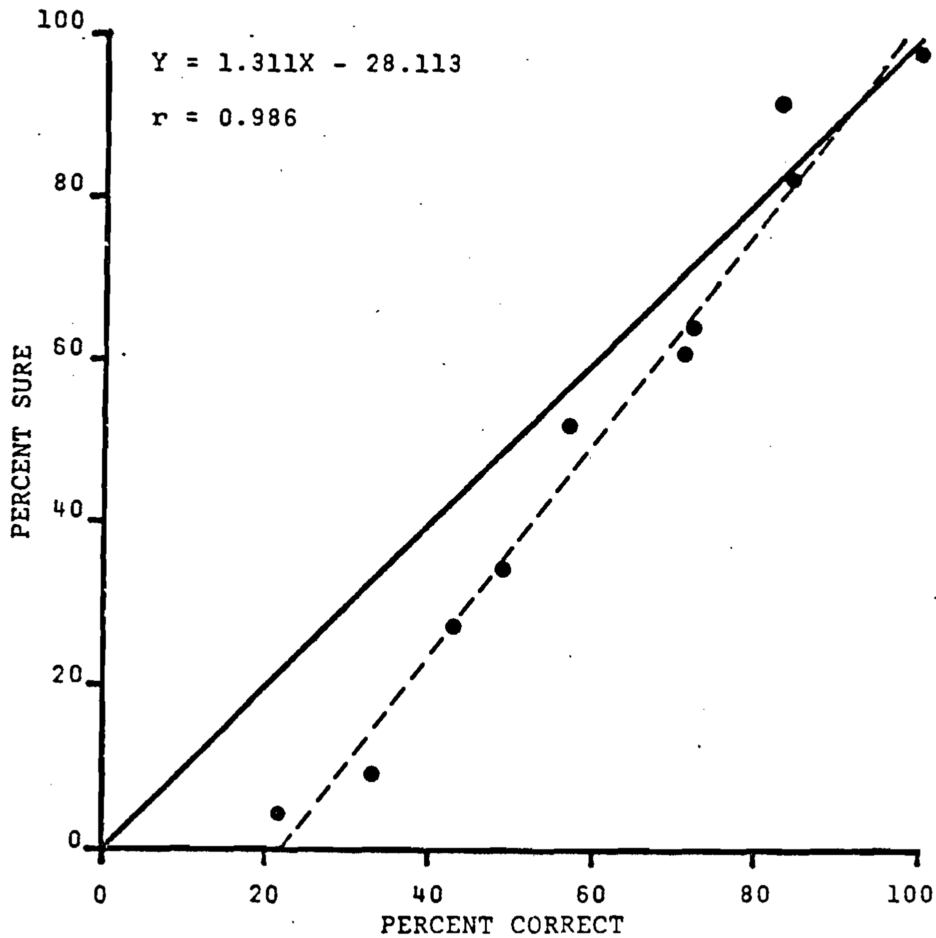


Figure 5(f). The mean percentage "Sure" responses as a function of the mean percentage "Correct" responses for Vocational students under Treatment KM.

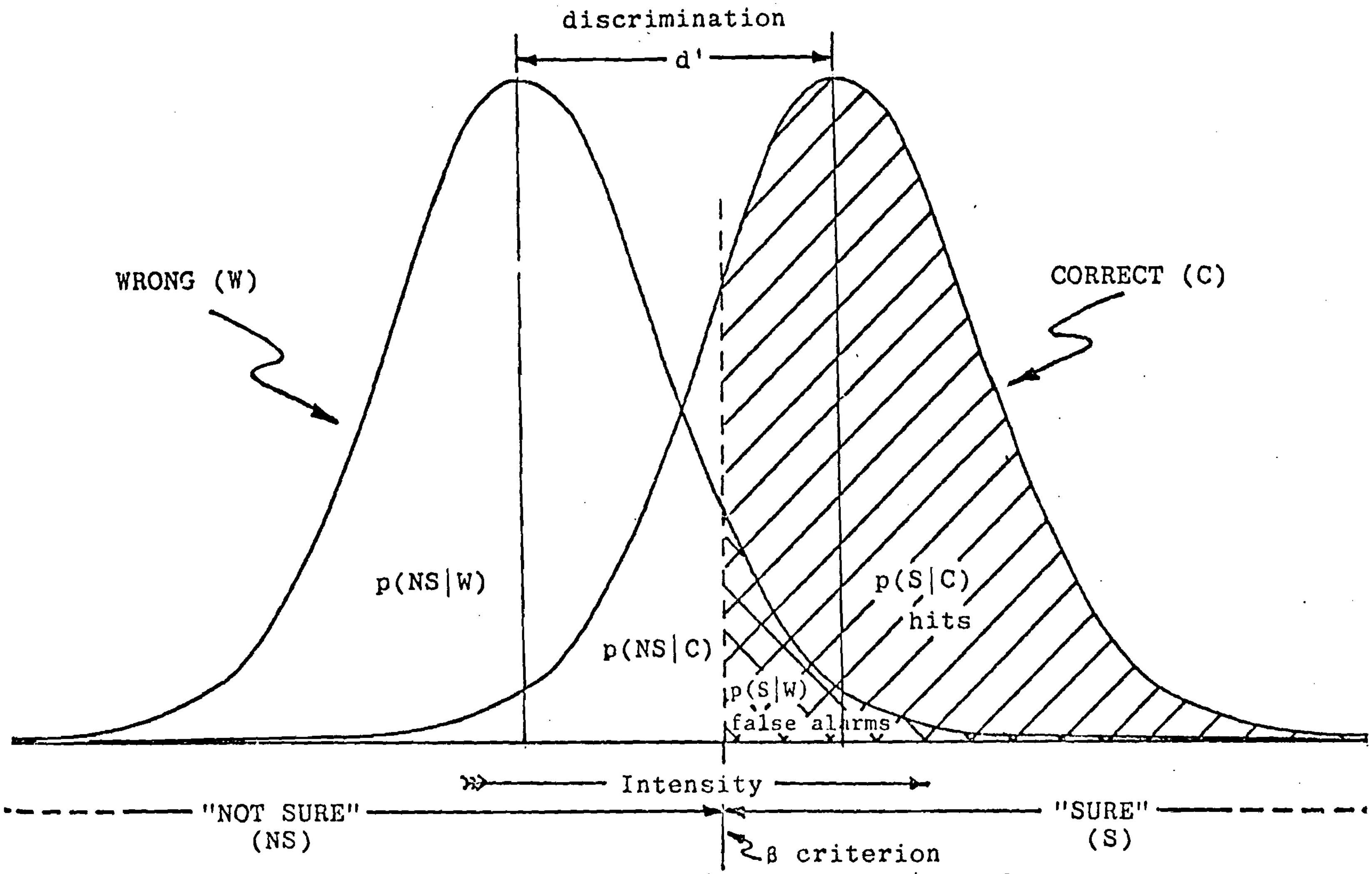


Fig. 6. Some concepts of signal detection theory related to the human self-assessment process,

The calculations (Green and Swets, 1966) utilized the hits, p(Sure|Correct), and false alarms, p(Sure|Wrong), to estimate the difference, d', between the means of the "noise" distribution (Wrong) and the "noise plus signal" distribution (Correct). Thus, the larger the d' measure then the more sensitive S is inferred to be in the detection of whether a selected response is correct or wrong. The mean values of p(S|C) and p(S|W) for treatments MK and KM are shown in Table IV for the first, middle and last third of the trials and for three groups of Ss. The associated d' values are presented in Table IV.

Inspection of Table IV reveals that generally the p(hit) and p(false alarm) is, say approximately 0.15, lower for treatment KM then for MK during the first third (0.56 vs. 0.44 and 0.39 vs. 0.24) and middle third (0.75 vs. 0.58) and 0.48 vs. 0.26) of the trials. This could be due to either (a) Ss under treatment KM are more conservative in saying that they are "Sure" of their answers so that they must be more confident that their answer is correct before they will call it "Sure" or (b) Ss under treatment MK may be better able to discriminate between correct and wrong responses because, say, some additional feedback concerning the correctness of their answer is produced by the execution of the M response. The mean d' values (0.56 vs. 0.64 and 0.93 vs. 0.99) in Table V corresponding with treatments MK and KM during the first third and middle third of the trials do not support the notion of better discrimination by Ss in treatment MK.

Table IV also shows that overall the mean p(hits) and p(false alarms) increases with practice for both treatments MK and KM. The apparent increase in d' from the first third of the trials to the last third suggests that the increased p(S|C) and p(S|W) values are at least partly due to an improved sensitivity of Ss to the correctness of their answers, i.e., their ability to discriminate correct from wrong answers improves with practice.

Table IV. The Mean p(Sure|Correct) and p(Sure|Wrong) Values for Treatments MK and KM for Each of the Three Groups of Ss Separately for the First Third, Middle Third, and Last Third of the Trials. Only data of Ss from which d' values could be determined are included; the number of individual values upon which each mean is based is shown in parentheses ( ).

		Treatment MK		Treatment KM	
Trials	Group	P(S|C)	P(S|W)	P(S|C)	P(S|W)
Last Third	PsF	0.85(4)	0.47	0.81(6)	0.60
	PsM	0.88(5)	0.50	0.82(3)	0.54
	VOC	0.89(4)	0.50	0.85(5)	0.40
	Mean	0.87	0.49	0.83	0.51
Middle Third	PsF	0.73(5)	0.43	0.50(7)	0.15
	PsM	0.77(7)	0.47	0.68(8)	0.32
	VOC	0.74(8)	0.53	0.56(9)	0.30
	Mean	0.75	0.48	0.58	0.26
First Third	PsF	0.44(9)	0.21	0.51(7)	0.28
	PsM	0.63(9)	0.45	0.50(9)	0.28
	VOC	0.60(7)	0.52	0.30(7)	0.17
	Mean	0.56	0.39	0.44	0.24

Table V. The Mean d' Values for Treatments MK and KM for the First Third, Middle Third and Last Third of the Trials for Each of the Three Groups of Subjects. Indeterminate values of d', for individuals Ss, i.e., P(hits) or P(false alarms) was equal to either zero or one, were not included; the number of individual values of d' upon which each mean is based is shown in parentheses ( ).

		Treatment	
Acquisition Level	Group	MK	KM
Last Third	PsF	1.18(4)	0.74(6)
	PsM	1.29(5)	0.81(3)
	VOC	1.33(4)	1.39(5)
	Mean	1.27	0.98
Middle Third	PsF	0.97(5)	1.06(7)
	PsM	0.89(7)	1.08(8)
	VOC	0.86(8)	0.82(9)
	Mean	0.93	0.99
First Third	PsF	0.86(10)	0.65(7)
	PsM	0.49(9)	0.68(9)
	VOC	0.33(7)	0.58(7)
	Mean	0.56	0.64

Internal Processing of the Self-Assessment Response.

A fairly extensive and speculative model of the human self-assessment process and some related processes has been outlined in the next section of this report. The processes of direct interest in this report are those involved in the selection (k) and the execution (K) of the self-assessment responses. In designing an experiment to test the effects of self-assessment on acquisition, it was desirable to include two treatments, MK and KM, which required Ss to make both an K and an M response but in different orders.

The present model proposes that the internal processing of the m and k responses is sequential in the order m then k; and the model implies that both responses are covertly available before either is executed. It may seem intuitively obvious that one must select the answer m response before one can assess its correctness. But Hart's (1967) work on the feeling-of-knowing and, to some extent, Nuttin's (1968) work on the feeling-of-success belie this obvious and simple conclusion. In any event, various logical possibilities by which answer and self-assessment responses may be selected, retrieved and executed are shown in Fig. 7.

It is supposed that differences in response latency between treatments MK and KM should reflect the order of the internal processing of the answer and self-assessment responses. Fig. 7 attempts to clarify some of the details involved in the latency of executing the K and M responses. The analysis depicted in Fig. 7 assumes that both the m and k responses are covertly selected before either the M or the K response is overtly executed; an alternative assumption will be considered later. Consider I-a in Fig. 7, which is the case in which the covert selection is in the order m then k, and the responses must be executed in the order M then K. The six steps involved following the stimulus onset are (1) select m, (2) select k, (3) retrieve the previously selected m, (4) execute M, (5) retrieve the previously selected k, and (6) execute K.

Now consider I-b, which is the case in which the covert selection is the same as above, i.e., in the

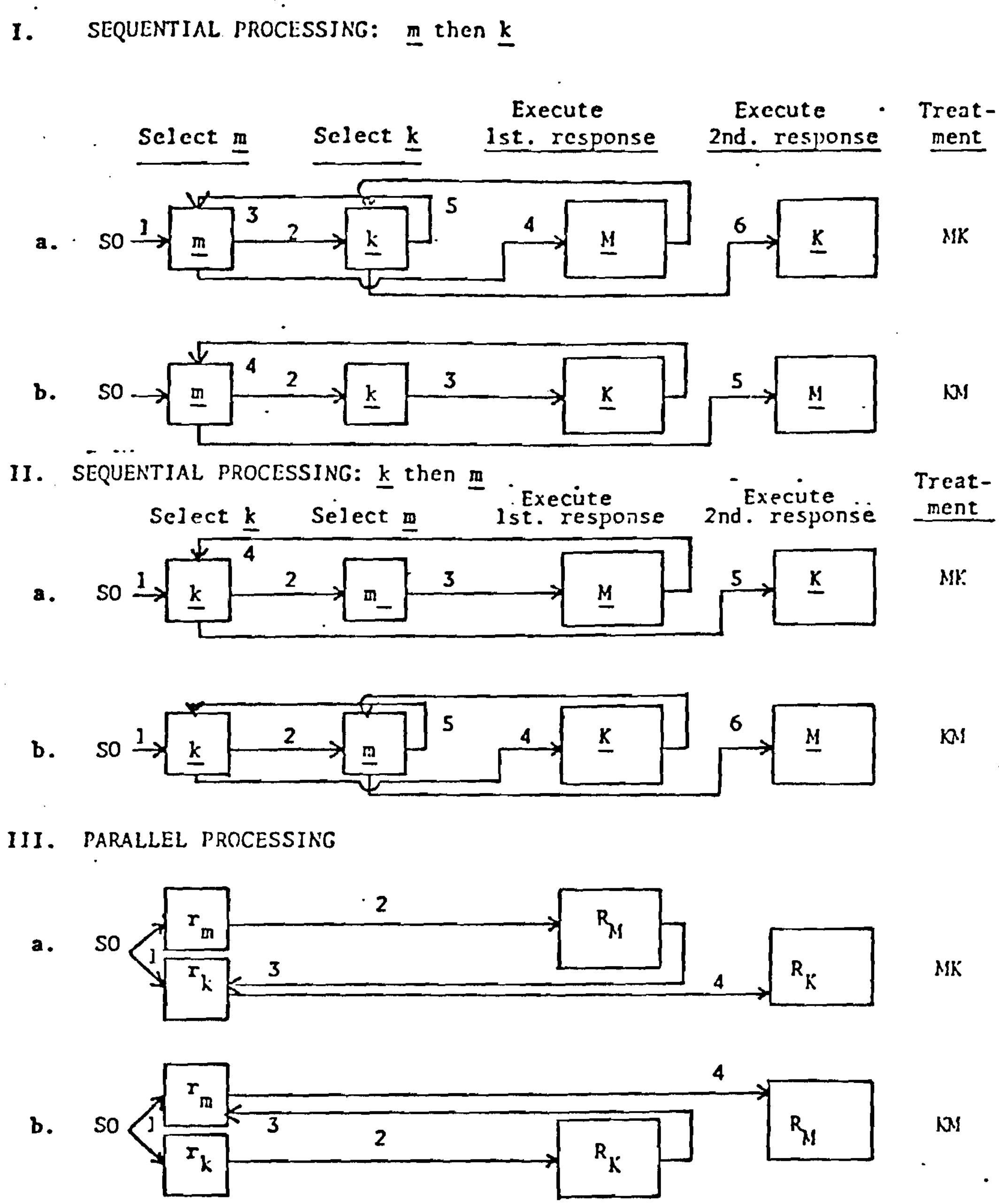


Figure 7. Various possibilities by which the answer and the self-assessment responses may be selected, retrieved and executed after the stimulus onset (SO); assumes that both the m and k responses are selected before either is executed.

order m then k, but the responses must be executed in the order K then M. No retrieval of the k response is required in Case I-b since k is executed as soon as it is selected. The other cases II-a,b and III-a,b may be similarly considered.

It seems reasonable to use the latency (call it LT2) from the onset of the stimulus to the initiation of the second response to obtain some data indicative of the order in which the m and k responses are covertly selected. If the m and k responses are processed in a parallel fashion then the LT2 should not be affected by the order in which the responses must be executed. If the internal processing is sequential in the order m-k, then LT2 should be longer for treatment MK than for KM because MK requires the retrieval of both the m and the k responses while treatment KM requires the retrieval of only the m response.

If the internal processing is sequential in the order k-m, then the LT2 should be shorter for treatment MK than for KM because treatment MK requires the retrieval of only the k response while treatment KM requires the retrieval of both the k and the m responses. In summary: if the internal processing is

a. sequential m then k, then $LT2_{MK} > LT2_{KM}$

b. sequential k then m, then $LT2_{MK} < LT2_{KM}$

c. parallel, then $LT2_{MK} = LT2_{KM}$.

The mean of LT2 values are shown in Table VI. An analysis of variance showed that the effect of treatments was significant, $F(3, 96) = 4.61$, $p < .01$, and that the effects of treatments did not significantly interact with either group or acquisition level. Calculating with mean combining the three groups and three acquisition levels shows that

$$LT2_{MK} = 3.544 > LT2_{KM} = 3.468$$

which suggests that the internal processing is either sequential in the order m-k (if the difference of 76 msec. is statistically reliable) or is parallel (if it is not); a test shows that this difference is not statistically reliable, $t < 1.0$, $p > .05$. The above analysis assumes that both the m and k responses are selected before either response is executed. If the

Table VI. Mean Latency (LT2) From The Onset of the Stimulus to the Initiation of the Second Response Under Treatments MK, KM, MX and XM for the Three Groups of Ss and at Three Different Acquisition Levels. Each mean is based on 10 Ss.

		Treatment			
Acquisition Level	Group	MK	KM	MX	XM
High	PsF	2.806	2.869	3.020	2.711
	PsM	3.155	2.776	3.003	2.540
	VOC	3.302	3.001	2.905	2.919
	Mean	3.088	2.882	2.976	2.723
Medium	PsF	3.407	3.447	3.524	3.199
	PsM	3.599	3.754	3.347	3.245
	VOC	3.650	3.791	3.443	3.561
	Mean	3.552	3.664	3.438	3.335
Low	PsF	3.936	3.863	3.856	3.527
	PsM	4.076	3.960	3.926	3.543
	VOC	3.968	3.753	3.654	3.686
	Mean	3.993	3.859	3.812	3.585
		3.544	3.468	3.409	3.214

M response is executed before the K response, then some modification of the covert k response might occur due to feedback to the person from the execution of M.

Also, the assumption that the m and k responses are completely processed internally before either is executed is only one possibility. An alternative assumption is that a response is executed as soon as it has been selected and is available. Fig. 8 shows some of the details of the response processes if this alternative assumption is made. Of immediate interest is the fact that this alternative assumption leads to inferences about the internal processing which are directly opposite to those inferences derived from Fig. 7.

For example, in Fig. 8 it may be seen that treatment KM should produce a longer LT2 than MK if the selection of the covert responses is sequential in the order m then k. For the comparable case in Fig. 7, which assumes that both responses are covertly selected before either is executed, the longer LT2 is associated with treatment MK rather than KM. Thus, it is necessary to determine which of the two assumptions is most tenable: (a) both the m and the k responses are covertly selected before either is executed or (b) a response, k or m, is executed as soon as it has been selected and the circumstances permits execution.

These assumptions may be tested by comparing treatments M and MK in terms of the latency (call it LT1) from the onset of the stimulus to the execution of the first, and M, response. If the k response is covertly selected prior to the initiation of the M response then the LT1 should be longer for treatment MK than for treatment M. Of course, if the m and the k responses are internally processed in a parallel fashion then the distinction between the two assumptions is of no interest.

The mean LT1 values are shown in Table VII separately for each treatment, group and acquisition level. Overall the LT1 for treatment MK is not longer than the LT1 for treatment M. Thus, in this task situation if subjects in treatment MK are selecting the k response prior to the execution of the

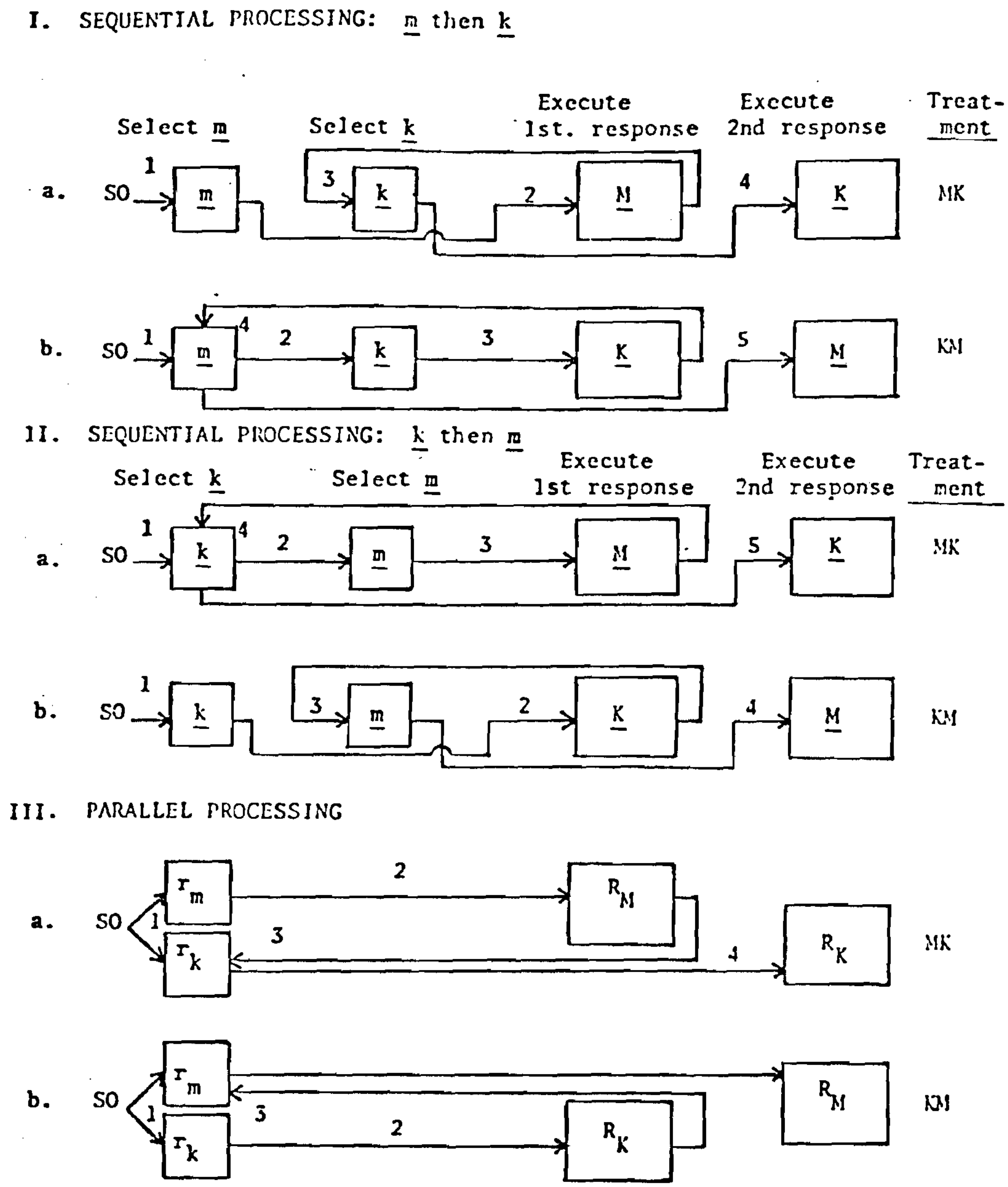


Figure 8. Various possibilities by which the answer and self-assessment responses may be selected, retrieved and executed; assumes that the required response, M or K is executed as soon as it is available.

Table VII. Mean Latency (LT1) From the Onset of the Stimulus to the Initiation of the First Response Button for the Three Groups of Ss and the Five Treatments at Each of Three Acquisition Levels. Each mean is based on 10 Ss.

		Treatments				
Acquisition Level	Group	M	MK	KM	MX	XM
High	PsF	2.555	2.090	1.914	2.428	1.448
	PsM	2.429	2.330	1.669	2.308	1.402
	VOC	2.323	2.476	1.968	2.268	1.715
	Mean	2.436	2.299	1.850	2.335	1.522
Medium	PsF	3.200	2.787	2.145	2.921	1.587
	PsM	2.965	2.825	2.287	2.624	1.634
	VOC	2.876	2.944	2.506	2.815	1.887
	Mean	3.014	2.852	2.313	2.787	1.703
Low	PsF	3.597	3.125	2.279	3.259	1.685
	PsM	3.015	3.262	2.468	3.241	1.780
	VOC	3.110	3.205	2.411	2.985	1.991
	Mean	3.241	3.197	2.386	3.162	1.819

M response, the processing of the *k* response is not reflected directly in the LT1 measure. It is possible that the subjects in treatment MK make their *M* responses quickly because they know that they must also make a *K* response within the 6 seconds time limit - while the subjects in treatment M need make only the *M* response within 6 seconds. Such a sense of urgency to make the *M* response might tend to obscure the reflection of the prior processing of the *k* response in LT1. There is some support for this notion of urgency by the observation that the subjects in treatment MX make their first, and *M*, response 100-200 msec. faster than do subjects in treatment M. From this point of view it may be noted in Table VII that the LT1 for treatment MK is 50-100 msec. longer than for treatment MX. This direction of difference in response latencies suggests that some kind of additional internal processing relative to the *k* response may be taking place in treatment MK.

A preliminary Model of the Human Self-assessment Process.

The model which is outlined in Fig. 9 borrows specific concepts and approaches from Kelley (1968), Adams (1971), Attneave (1974) and Miller, Galanter & Pribram (1960). The model is intended to provide a detailed framework for considering the self-assessment process and is not intended necessarily to portray the underlying physiological mechanisms involved. Generally, the capital letters indicate events which are observable such as an overt response of the person and the small letters indicate implicit or covert responses, events or states.

Each block will be considered in turn and the rationale for their inclusion will be discussed:

1. The goal. It is assumed that the responses or outputs of the person are selected and executed for the purpose of attaining certain desired goals at any moment in time. Kelley (1968) adopts this position essentially when he points out that the typical feature of the living organism is the conception and choice among goals (p. viii). The notion that organisms behave in accordance with purposes is a reasonable assumption (Miller, Galanter & Pribram, 1960). And Nuttin and Greenwald (1968) state, "...the outcome of an action is regarded as playing a fundamental role in behavioral processes. Specifically, a future outcome can be said to determine behavior in the sense that the outcome is 'intended' prior to the performance of the action and the anticipation of the outcome subjectively appears to have the power of eliciting the action (p. 2)". In the relatively simple paired-associates task employed in the experiment, it is assumed that S's "goal" to be correct.

2. The Internal Model. This is the process by which the person is able to implicitly predict the possible consequences or outcomes of various implicit responses which he may wish to test in fast-time. Attneave (1974) diagrams it in "a somewhat oversimplified way (p. 494)" as:

$$S_1, R \rightarrow S_2$$

and notes, "If situation S_1 obtains at a given time, and I do R, then situation S_2 results. If I know this, I know how to change situation S_1 into situation S_2 (p. 494)." It is assumed that an individual possesses a

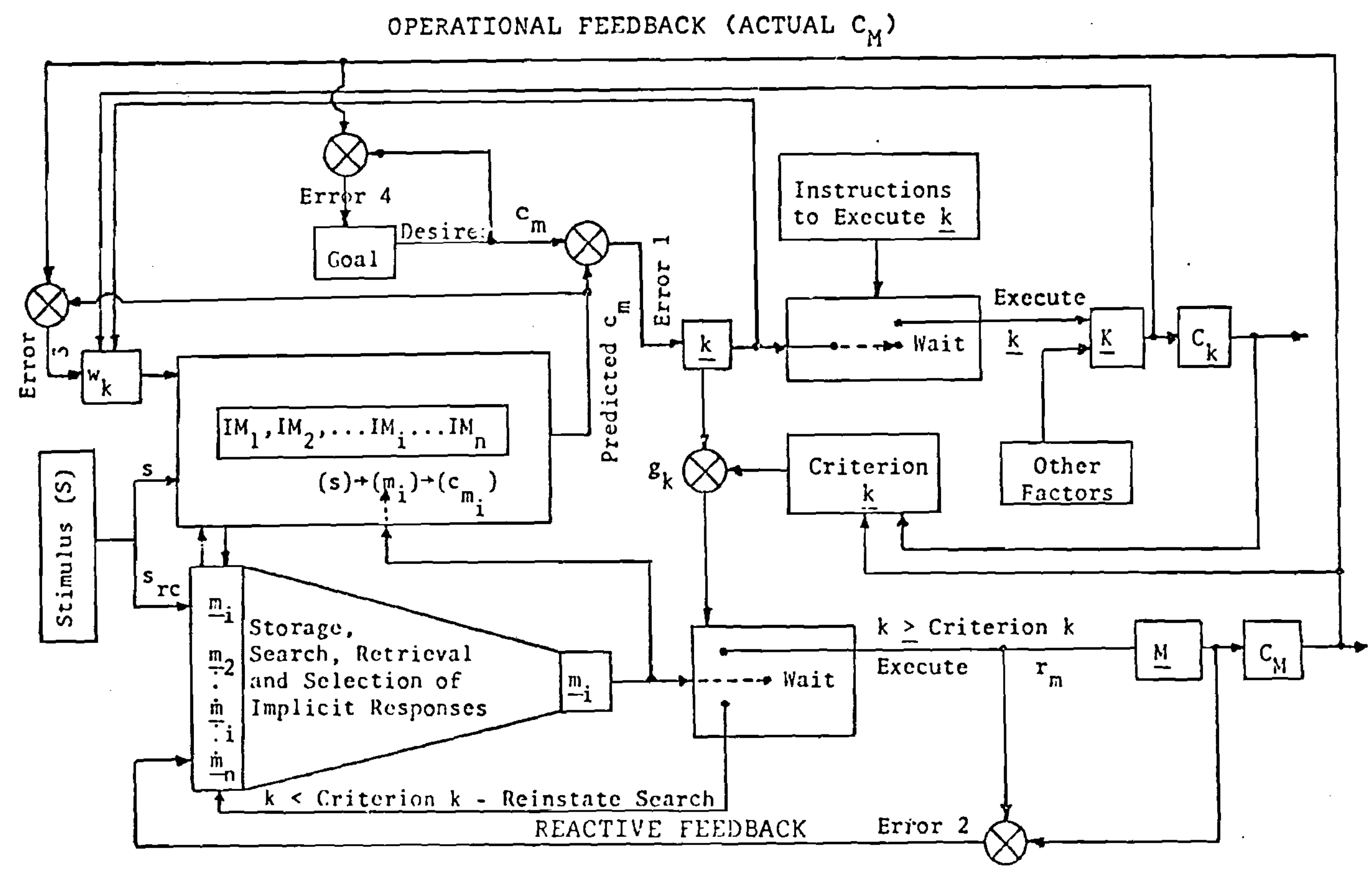


Figure 9. A preliminary model of the human self-assessment process

fairly extensive repertoire of such internal models, from which he selects one (or some) of the models depending upon the situation which exists and the goals which he seeks. Presumably both the repertoire of models and the specifics of each internal model are built up through learning and experience in which the consequences of both covert and overt responses are predicted by the individual and then compared with the consequences which are actually produced later by the response which is selected and executed. This proposed view seems consistent with Levine's (1975) characterization of adult human learning as the testing of hypotheses in a situation and learning by searching for and finding the correct rule.

Others (Miller et al. 1960) have used the term, "Image," to describe, "... all the accumulated, organized knowledge that the organism has about itself and its world ... (and) it includes ... his values as well as facts ... (p. 17)." Also, Deese (1969) seems to imply a similar internal process which he calls "understanding" when he states, "understanding only signals the potential for appropriate imagery, linguistic operations and other cognitive activity. The operations may not occur unless there is some challenge, either self-induced or external, to prompt them (p. 516)." However, he does not clearly distinguish between (a) the existence of the potential, which would be comparable to an Internal Model and (b) the recognition by an individual that he possesses the potential, which would be comparable to what I have called self-assessment. On the one hand, he states, "Each human being is capable of recognizing a state of understanding (p. 516)," which suggests that the state of understanding is separate from an individuals recognition of it. On the other hand, he states, "Understanding is the inward sign of the potential for reacting appropriately to what we see and hear (p. 516)," which suggests that the concept of understanding includes the recognition by an individual that he possesses some potential. Within the context of my proposed model this is an important distinction.

3. <u>Stimulus</u> (S). This is intended to be the explicit stimulus which the experimenter presents on a trial, as well as those features which define what Attneave calls "S_1" or situation 1. This "Stimulus" performs two functions in the model:

a. It provides the input data to the person so that he can describe, to whatever extent he can or will, the situation 1 which prevails at the time. This is labelled "s" in Fig. 9.

b. It serves as a retrieval cue which initiates the memory search and retrieves or selects an implicit response. This selected response, $\underline{m}_i$, serves as an input to the (s-$\underline{m}_i$- c) Internal Model of the person which permits him to implicitly assess the possible consequences as specified by the goal. This retrieval cue property of S is labelled "s_{rc}" in Fig. 9. The intent is to emphasize the importance of the retrieval cue in the retrieval process as distinct from storage, memory and forgetting. The distinction is of general importance in recall and recognition (Tulving, 1974), but it is of especial importance in considering the human self-assessment process. The fundamental interest in the self-assessment process, at the present time, is with the relationship between (a) information which may, or may not, be stored in a person's memory and (b) the person's ability to validly and reliably determine that it is, or is not, stored in his memory. The person's demonstration by recall or recognition requires, in addition to its being stored, that he retrieve it under the circumstances which exist at the time of an inquiry or a demand that he utilize in some fashion the knowledge which is stored. In any case, I am not merely making the general and traditional distinction between learning, on the one hand, and performance, on the other.

4. Covert self-assessment response ($\underline{k}$). The anticipated consequences of the selected $\underline{m}_i$ are predicted by means of the Internal Model and this prediction is the output of the Internal Model. As Broadbent (1973) points out, "... there is reasonable ground for believing that our brains calculate upon a model of the world the various consequences that will arise from different actions (p. 180)." The anticipated or predicted consequences, c_{m_i}, are compared with the desired

consequences as specified by the "Goal." The discrepancy is labelled "Error 1". It is assumed to be inversely related to _k_, i.e., the greater the discrepancy then the lower the implicit confidence. This _k_ may be made observable by simply asking the person - presumably with some distortions in the translation due to a number of other factors, however.

The _k_ is assumed to serve two purposes:

a. It serves indirectly as a "gatekeeper" for the response m_i in the following way. The covert m_i will not be executed unless _k_ attains some criterion level, as indicated by the comparison of _k_ and the "criterion _k_." It is further assumed that, once the criterion _k_ is attained, the vigor, speed, etc. of _M_ is directly related to _k_; the greater is _k_ then the more vigorously, quicker, etc. the _M_ response is executed.

b. Also, it serves as a weighting factor in the operational feedback look from the consequences, C_M, of _M_ back to the Internal Model, as will be described later.

Regarding the gatekeeping function of _k_, you will note in Fig. 9 that when _m_ is produced then any one, but only one, of three things can happen: (1) _M_ can be executed, (2) a search of another m_i can be initiated or (3) m_i can be held in abeyance until either 1 or 2 is selected. It is assumed that the latency of the reinitiation of a search is also related to the value of _k_, such that the more confident the person is that the selected m_i is _not_ the appropriate response, the quicker the search is re-initiated.

The findings of Murdock & Dufty (1972) are consistent with this assumption. They found that the latency with which a visually presented item was recognized as having been a member of a previously presented list or as having not been a member of the list was inversely related to the confidence expressed (on a 6-point scale) by the subjects.

Murdock & Dufty generally interpret this finding, and others, as being consistent with the notion that the speed and confidence with which an item is recognized as being or not being a member of a previous list depends fundamentally upon the strength of the underlying memory trace rather than involving any separate "self-assessment" process.

However, Bernbach (1967) points out that a strength theory predicts that certain features of the receiver-operating-characteristic curves (which may be produced by a signal detection analysis of some learning-confidence data) should be related to factors (such as the serial position of an item) which influence the strength of a response. He presents evidence which fails to support this prediction. Thus, Bernbach's conclusions are consistent with the view that some process(es) separate from the learning process is (are) operating to produce confidence responses. The model of the self-assessment process outlined in Fig. 9 speculates as to the manner in which this separate process may function.

A situation which presents some difficulties for the proposed model relative to _k_ is one in which the response made is either "correct" or "wrong" as in the paired-associates learning task used in the present experiment. The difficulty arises because there does not seem to be various degrees of discrepancy between the "goal" and the "predicted consequences." But the view that the response in a paired associates learning task is either "correct" or "wrong" may obscure some relevant circumstances. For example, for _S_ to make a "correct" response he must correctly accomplish a number of component sub-tasks or activities, e.g., he must detect and identify the stimulus, attend to the attempt of searching his memory for a response, execute the response within the time limits. If any one of these or many other components is deficient then the response is labelled "wrong." Thus, _S_ could correctly accomplish 90% of the components and still fail to make a "correct" response." It is suggested that, under normal circumstances, the consequences predicted by the internal model could reflect the accumulation of all of the components. If _S_'s goal is to make correct responses 100% of the time, then the notion that _k_ is inversely related to the discrepancy still seems reasonable.

In a paired associates task, the $\underline{S}$ is informed somehow as to the correctness of his response. For example, after each response the $\underline{S}$ may simply be told "correct" or "wrong"; or the stimulus along with the correct response may be presented so that $\underline{S}$ is directly informed as to the correct response and is indirectly informed as whether his response was correct by comparing the response that he made with the correct response. In normal experimental situations $\underline{S}$ will usually repeat a response if it has previously been followed by "correct" and not repeat a response if it has been followed by "wrong."

Buchwald (1969) and others (d'Ydewalle & Eeleen 1975) have proposed that the repetition or non-repetition of such a response which has been previously made depends upon whether the person (a) recalls the previous response and (b) recalls the outcome or feedback information relative to the previous response.

In the Internal Model (Fig. 9), these two recollections presumably would refer to (a) the retrieval of the response $\underline{m}_i$ when situation $\underline{s}$ is presented and (b) the ability to predict the consequences, c_{m_i} if $\underline{m}_i$ were to be made from situation $\underline{s}$ exists.

From this point of view, the confidence $\underline{k}$ would be a function of

a. the probability that $\underline{m}_i$ will be retrieved (and tested) when situation $\underline{s}$ (stimulus) is presented, which is equivalent to the probability of recalling the response which was previously made to the stimulus

b. and the probability that $c_{\underline{m}_i}$ will be recalled when m_i is tested in the Internal Model, which is equivalent to $p(c_{\underline{m}_i}|\underline{m}_i)$ or the probability of recalling the outcome.

For illustration, say that only two responses, $\underline{m}_i$ and $\underline{m}_2$, are in the response repertoire when situation $\underline{s}$ is presented; and that

$$p(\underline{m}_1|s) = 0.8$$

and

$$p(\underline{m}_2|s) = 0.2$$

Also, assume that the learner recalls that, when $\underline{M}_1$ has been made in the past, it has been called "correct" 80% of the time and called "wrong" 20% of the time. For $\underline{M}_2$ assume that the learner recalls that it has been called "correct" 20% of the time and "wrong" 80% of the time.

For our purposes this, in fact, may be a suitable description of a two-light prediction task. In a two-light prediction task in which p(light 1) = 0.8 and p(light 2) = 0.2, the relative frequency of the person's choice of light 1 and light 2, under "no-payoff" conditions, is typically found to be approximately 80% and 20%, respectively--a matching choice strategy (Hurvich, 1970). If the experimental situation is altered so that the person receives a reward for making a "correct" prediction and a "punishment" for making a "wrong" prediction, then usually the person will tend toward a pure strategy of predicting the most frequently occurring event all of the time.

It should be noted, however, that the <u>expected proportion of correct predictions for $\underline{M}_1$ is 0.8</u> regardless of whether the person employs a matching or a pure strategy, i.e., approximately 80% of the $\underline{M}_1$ responses will be correct. Similarly the expected proportion of correct predictions for $\underline{M}_2$ is 0.2 regardless of how the person distributes his responses between $\underline{M}_1$ and $\underline{M}_2$ (provided $\underline{M}_1$ and $\underline{M}_2$ is chosen at all).

As was stated earlier, assume that the learner's "Goal" is to give correct answers with certainty or as near to certainty as he can. The assumption is speculative, but is consistent with Sampson and Chen (1971) in their proposal model of human binary prediction behavior. Presumably the person retrieves some response, say $\underline{m}_1$, and tests it through the Internal Model:

$$(s) \rightarrow (m_1) \xrightarrow{0.8} \text{"Correct"} \ (c_{m_1})$$

If some goal is to be "correct 100% of the time" and the consequence predicted by $\underline{S}$ is that he will be "correct 80% of the time", then the discrepancy is "20%." Now the present model assumes that the confidence, $\underline{k}$, is inversely related to the discrepancy produced by a comparison of the "desired outcome" and the "predicted outcome." So, for purposes of consistency of the model one may view the relationship between the discrepancy and $\underline{k}$ as shown in Fig. 10.

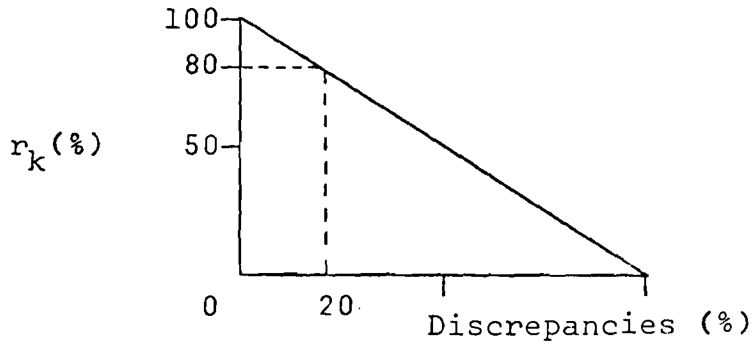


Fig. 10. Discrepancy (Desired c_m-Predicted c_m).

For example, when the "predicted c_m" is 80%, then the discrepancy is 100% minus 80%, which is 20%; and the confidence in the correctness of the response is 80%.

This line of thought makes it explicit that k may depend upon the probabilistic relationships between m_i and c_{m_i} in the person's Internal Model. There are at least two major sources of this $m \rightarrow c_m$ uncertainty. First, the uncertainty could be due to the incomplete learning of the $[(m \mid s_1) \rightarrow (c_m)]$ relation by the person; this might be called "model uncertainty." Second, the uncertainty could be inherent in the real-world situation which the Internal Model represents; this might be called "real-world uncertainty."

The two-light prediction task discussed earlier, in which the outcomes are probabilistically related to the responses, is an example of "real-world uncertainty." It is expected that the amount of real-world uncertainty determines the limit of the confidence, k , which the person may attain. In a two-light prediction task in which the p(Light 1) is 0.8, the maximum confidence S can attain for his (prediction) choice of Light 1 is 80% because the "real-world uncertainty" is at that level. On the other hand, in a typical paired-associate learning task, the $[(s_1) \rightarrow (m_1) \rightarrow (\text{correct})]$ relationship is fixed and, thus, the real-world uncertainty is virtually 0; and S can reasonably be expected to attain a 100% confidence when the internal models is appropriately developed. In a two-light prediction task a person can be influenced to depart from a "matching" choice behavior toward a "pure" choice behavior by increasing the payoffs for being correct. However, the person's confidence in the correctness of his predictions would not be expected to increase in a comparable fashion. This is the case because the expected proportion of choices of one light (or the other) which is "correct" is independent of the number of times the light is chosen.

Thus, the confidence which a person possesses regarding the correctness of his anticipated response may depend upon both (a) the discrepancy between the goal and the predicted c_m, and (b) the probabilistic relations between $m_i \rightarrow c_m$ which exists in the person's Internal Model[1].

5. Response repertoire. At a molecular response level, this is quite similar to Adams' (1971) concept of a memory trace distribution, i.e., at a particular moment in time there is a variety of (simple) responses available with associated probabilities of being selected. At a more molar response level, one may think of the repertoire as being composed of a number of alternative elaborate sequences-of-responses (or possible plans of action) which might be executed by the person (Miller et al, 1960; Attneave, 1974).

In Adams' theory of motor learning, two separate trace distributions are necessary largely because, "If the agent that fires the response also is the reference against which the response is tested for correctness, the response must necessarily be judged correct because it is compared against itself (p. 125)." However, the theoretical demand for a separate memory trace and a perceptual trace distribution is eliminated in the proposal model of the human self-assessment process. In the proposed model, the comparison which determines the self-assessment of the correctness of a potential response is between (a) the predicted consequences, c_m, of a response, m, and (b) the desired consequences as indicated by the goal.

One fundamental question is related to the heirarchical level of the response being considered; a second fundamental question is concerned with the location of the feedback or control loop and the locus of the standard against which the appropriateness of a response is assessed. Roy and Martenuik (1974) point out that feedback information relative to learning a simple motor response may be provided by extrinsic operational knowledge of results, intrinsic response-produced stimuli, or central monitoring of the efferent outflow of a response. A fourth (cognitive) source of feedback information is suggested which is the anticipated consequences if some response were to be executed. Furthermore, it is reasonable to suppose that the location of the feedback loop (extrinsic, intrinsic, central, or cognitive) as well as the locus of the standard may depend upon the heirarchical level of the response. For example, there is some evidence (Roy & Martenuik, 1974) that simple motor responses of, say, less than 150 msec. are controlled by different loops than are similar response of 1 sec. or longer.

In any event, the function of selecting and initiating a response is carried out in the proposed model in the following way:

a. An implicit response, m_i, is tentatively retrieved from the response repertoire for testing through the Internal Model.

b. The overt execution of a response, M depends upon the level of k which is produced by the "Error 1." If the k is greater than the criterion k, then the M is initiated; otherwise a new m_i is retrieved for testing, etc. The problems associated with the person's stopping his search is related to the self-assessment process. But very little seems to be known about the factors involved in the cessation or the continuation of the memory search. For the sake of tentative theoretical closure, it will be assumed that the reiterative retrieval process continues until the criterion k is reached, the response repertoire is exhausted, or until there is a readjustment of the criterion k or goal.

There is one more feature of the response repertoire which needs to be mentioned. Once the implicit response, m_i, is initiated through g_k, then its overt execution, M, is monitored through the reactive feedback loop in an attempt to maintain a suitable fidelity between m and M. One could, in the model, provide for this function by having "Error 2" enter the process immediately prior to M but after g_k or even immediately prior to g_k. However, having "Error 2" enter the response repertoire indicates that the M which is executed also modifies the response repertoire. Whether the "specification" of m which serves essentially as the criterior for the execution of M whould be before or after the gatekeeper, g_k, is not clear. As research is accomplished this question, along with others concerning the model, might be clarified.

Another assumption of the proposed model which is of interest, but not of compelling attention at the moment, is that there may be stored in (long term) memory an extensive repertoire of Internal Models of an $[(s) \rightarrow (m) \rightarrow (c)]$ kind. Presumably a person retrieves a specific Internal Model based upon his analysis, (s), of the Situation, (S). It is this specific Internal Model into which a selected response, $\underline{m}_i$, is inserted to anticipate the consequences, c_m, of that response. Thus, one's confidence or self-assessment could be in error or misleading because an inappropriate Internal Model is used for anticipation of the consequences of a response. But this assumption involves behavior at a more molar level and more complex than seems to be necessary to consider at the moment.

6. <u>Operational feedback</u>. The overt response, $\underline{M}$ produces certain consequences, C_M, in the real-world. This is what Attneave (1974) calls, "Situation 2." Information concerning the actual consequences are fed back and compared with the predicted consequences, c_m, of the response. Any discrepancy between C_M and c_m, which is called "Error 3" in Fig. 4, may result in a compensatory modification of the Internal Model; this is part of what may be called learning or increasing ones knowledge.

The covert and overt self-assessment responses, $\underline{k}$ and $\underline{K}$, are hypothesized to play an important role in the operational feedback loop. They serve to "weight" the Error 3. It is hypothesized that, for a given size of discrepancy between the predicted and actual consequences of a response, the extent to which the Internal Model will be modified is inversely related to the confidence of the person, i.e., the greater the confidence then the less the modification of the Internal Model.

This hypothesis is based on the notion that, generally, a strong belief is not likely to be modified as a consequence of a single disconfirmation. It is tentatively presumed that the less strong the belief, i.e., the lower the confidence in the correctness of the response, then the more the Internal Model is subject to being modified by a disconfirmation.

The notion of a misinformed, in contrast to an uninformed, individual is relevant here (Shuford, 1967). The difference between a misinformed and uninformed individual may be represented as shown below:

	Confidence Expressed	
	0%	100%
Response Correct	Uninformed(?)	Informed
Response Wrong	Uninformed	Misinformed

Thus, a misinformed state is indicated by a high confidence in the correctness of a response which, in fact, is wrong. In terms of the Internal Model, a misinformed state would suggest that the association between $(R|S_1)$ and (S_2) is well established but wrong, i.e., the implicit prediction, c_m, based on "s" and "$\underline{m}$" is wrong.

If (a) a misinformed state of knowledge is reflected by a high confidence in a wrong or inappropriate response and (b) if, in fact, the extent to which the feedback produces a modification of the Internal Model is influenced by the confidence as described above, then the instructional or training problem is somewhat different for the misinformed, than it is for the uninformed individual.

Also, C_M may directly influence the criterion confidence, $\underline{k}$, on subsequent trials. Thus, the actual consequences of making a response, $\underline{M}$, may modify the level of confidence required to execute the response on the next and subsequent trials. The most straightforward way of incorporating the effect of C_M on the criterion $\underline{k}$ is simply to consider the effect as a shift of the criterion within the framework of signal detection theory (Green & Swets, 1966).

7. <u>Consequences of overt self-assessment response</u> (C_K). In most human learning or performance research, no overt self-assessment response is required of subjects. In the relatively few studies in which $\underline{K}$ has been required, the consequences, C_K, have not been experimentally manipulated. Based upon the preliminary model of the self-assessment process outlined here, it is explicitly hypothesized that the C_K will affect certain characteristics of human performance, $\underline{M}$, presumably by affecting the criterion $\underline{k}$ of the gatekeeper, g_k.

8. Goal modification. Central to Kelley's (1968) discussion of the human as a component in the control process is the notion that people are able to conceptualize possible goals and to choose from among them; and he points out that this most important process is the least understood. I wish not to confront this problem area, but merely to point out that, given that some goal exists at one moment in time, it is subject to being modified. These modifications of a goal depend upon a variety of variables such as instruction, constraints, changing abilities of the organism, etc. The only factor which is depicted in Fig. 9 as affecting the Goal is "Error 4," which is the result of a comparison by the individual of the desired consequences with the actual consequences or perhaps more precisely, he compares his conception of the desired consequences with his perception of the actual consequences. As a result of this comparison he may modify (raise or lower) his goals.

Miller et al (1960) seem to view the goal modification in a similar way when they say, "An alternative to the stop-rule (for searching) is a modification of the conditions that are imposed in the test phase. After searching unsuccessfully for a pen, we settle for a pencil (p. 171)." Their view makes it explicit that the goal may be modified prior to the overt execution of a response as well as after the consequences of an executed response have been compared with the desired consequences.

Comment. A fairly extensive and highly speculative model of the human self-assessment process and other processes related to it has been outlined. It should be emphasized that the problem area of interest here is the human self-assessment process and its relation to learning. For example, the following processes are only of incidental interest: Goal definition, memory search and retrieval, response selection and execution, the consequences of responding, and most of the feedback information. The component processes which are of primary interest are the implicit and explicit responses ($\underline{k}$ and $\underline{K}$) concerning the confidence of the person; it is also of central interest to determine the manner and extent to which the confidence of the performer or responder affects human learning and performance. It was felt necessary to speculate, by constructing a preliminary model, about the kinds of interactions between the self-assessment process (which result in some level of confidence) and the other human processes.

References

Adams, J.A. A closed-loop theory of motor learning. Journal of Motor Learning, 1971, Vol. 3, No.2, 111-149.

Anon. Intermedial Raytheon Integrated Learning Systems, Raytheon Corporation, Englewood, New Jersey, .

Attneave, F. How do you know? American Psychologist, 1974, Vol. 29, No. 7, 493-499.

Bernbach, H.A. Decision processes in memory. Psychological Review, 1964, 74, 462-480.

Broadbent, D.E. In defence of empirical psychology. Great Britain: Methuen & Co. Ltd., 1973.

Buchwald, A.M. Effects of "right" and "wrong" on subsequent behavior: a new interpretation. Psychological Review, 1969, 76, No. 2, 132-143.

Carpenter, C.R., Eggleton, R.C., John, F.T. and Cannon, J.R., Jr. The classroom communicator. Technical Report SDC-269-7-14, Human Engineering Branch, Special Devices Center, Port Washington, L.I., N.Y., 1950.

Deese, J. Behavior and fact. American Psychologist, 1969, 24, 515-522.

d'Ydewalle, G. and Eelen, P. Repetition and recall of "right" and "wrong" responses in incidental and intentional learning. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, No. 4, 429-441.

Green, D.M. and Swets, J.A. Signal detection theory and psychophysics. N.Y.: John Wiley and Sons, Inc., 1966.

Hart, J.T. Recall, recognition, and the memory-monitoring process. Stanford University Dissertation, 1966.

Kelley, C.R. Manual and automatic control: A theory of manual control and its application to manual and to automatic systems. New York: John Wiley and Sons, Inc., 1968.

Levine, Marvin. A cognitive theory of learning: Research on hypothesis testing. New York: John Wiley and Sons, 1975.

Levine, J.M., Samet, M.G. and Brahlek, R.E. Information seeking with input pacing and multiple decision opportunities. Human Factors, 1974, Vol. 16, 384-394.

Lichtenstein, S.C. and Fischhoff, B. Do those who know more also know more about how much they know? Oregon Research Institute Bulletin, Vol. 16, No. 1, Feb. 1976, Eugene, Oregon 97403, PO Box 3196.

Miller, G.A., Galanter, E. and Pribram, K.H. Plans and structure of behavior. New York: Holt, Rinehart and Winston, Inc., 1960.

Miles, J. Eliminating the guessing factor in the multiple choice test. Educational and Psychological Measurement, 1973, Vol. 33, 637-651.

Murdock, B.B., Jr. and Dufty, P.O. Strength theory and recognition memory. Journal of Experimental Psychology, 1972, 94, No. 3, 284-290.

Nuttin, J. and Greenwald, A.G. Reward and punishment in human learning. London: Academic Press Inc. Ltd., 1968.

Roy, E.A. and Martenicuk, R.G. Mechanisms of control in motor performance: Closed-loop versus motor programming control. Journal of Experimental Psychology, 1974, 103, 985-991.

Sampson, J.R. and Chen, I.N. Computer simulation of multiple strategies in human binary prediction. Psychological Reports, 1971, 29, 1207-1221.

Shuford, E.H. and Massengill, H.E. The relative effectiveness of five instructional strategies. Interim Technical Report, Contract No. AF 49(638)-1744, ARPA Order No. 833, 1967.

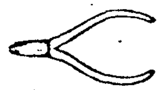
Tulving, E. Cue-dependent forgetting. American Scientist, 1974, 62, 74-82.

Wade, T.C. Relative effects on performance and motivation of self-monitoring correct and incorrect responses. Journal of Experimental Psychology, 1974, 103, No. 2, 245-248.

Woodworth, R.S. and Schlosberg, H. Experimental Psychology. New York: Holt, Rinehart and Winston, 1954.

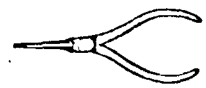
APPENDIX A. Stimuli and Responses Used in the Primary Learning Task.

SPLIT



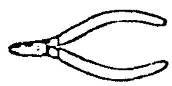
CLIP

SHAPE



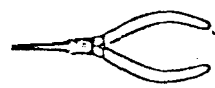
BEND

CUT

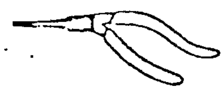


SNIP

TWIST



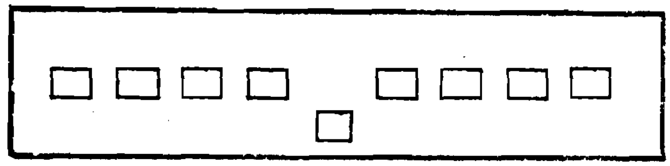
FORM



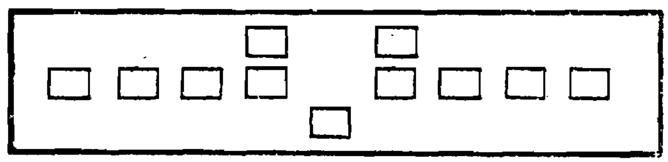
Appendix B. Button Arrangements on the Panel for the Five Different Treatment Conditions

<u>Treatment</u>

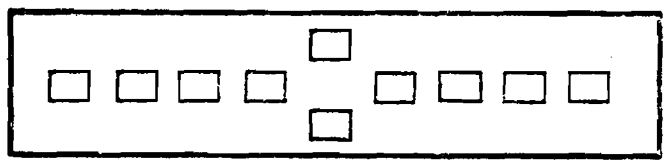
M



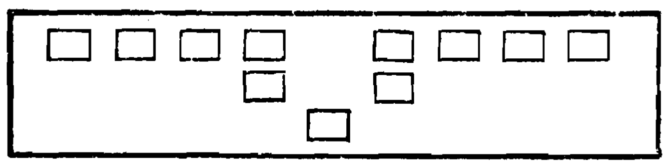
MK



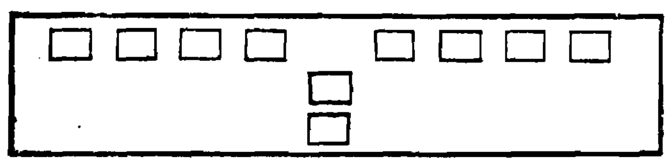
MX



KM



XM



Appendix C. The Two Button Orders Used.

Button Position		Order 1	Order 2
(Left)	1	TWIST	FORM
	2	CLIP	SPLIT
	3	FORM	SHAPE
	4	SPLIT	SNIP
	5	SHAPE	CUT
	6	SNIP	BEND
	7	BEND	CLIP
(Right)	8	CUT	TWIST

Appendix D. Instructions to Subjects

Group R_M

Please have a seat in this chair. (POINT & LET S GET SETTLED) In this experiment, your task is to learn to identify each of 8 different pliers by name. A picture of a pair of pliers will appear on this screen. (POINT) You simply press a button in this answer row to indicate the name of the pliers. (POINT SWEEPINGLY) You should try to press the button as quickly and as accurately as you can.

When we start the experiment, you will put on these earphones. (HOLD HEADSET UP) During the experiment, each trial will take place in the following order: First, you must hold this "start" button down. (POINT & PRESS) You will hear a short tone through the earphones which will inform you that the pliers is about to be presented. After the tone, the pliers will appear on the screen and stay on for 6 seconds. During the 6 seconds, you have to release the start button, then press the answer button as fast and as accurately as you can. Shortly after you have pressed the answer button, the picture will go off; and, then the same pliers will appear again on the screen together with the correct answer for 4 seconds so that you can see whether your answer was correct.

Shortly after, the brief tone will sound again to signal that the next pliers will be presented. You should make sure that you are pressing the "start" button when you hear the tone and that you hold it down until the slide is presented. If you don't do this, a second tone will be presented, then if you don't press the "start" button immediately, a wrong answer will be recorded and you will be shown only the pliers and the correct answer after a 7 second period with nothing on the screen.

When the experiment first begins, the 8 pictures of the pliers, along with their correct names will be presented one time for 5 seconds each so that you can study them. From then on, the pictures will appear alone and you will have to name them by pressing the proper buttons.

The 8 pictures of pliers will be repeated in a different order each time, until you have gone through the sequence twice in a row, without an error. The start of each sequence will always be indicated by a dot on the screen

It will help you to tell the difference between the pliers if you notice that there are 2 different heads, 2 different shapes of handles and the handles may be plain or cushioned. (SHOW THREE PICTURES OF PLIERS FOR 45 SECONDS, POINTING OUT TO S THE CRITICAL CUES--THREE 8 x 10's THAT CONTAIN ALL FEATURES--RANDOMIZATION LISTED ON SCHEDULE)

Once again: the 8 pliers along with their correct names will be presented first and you should study them. Then your task begins with a dot on the screen and your finger on the "start" button. The brief tone sounds, followed by a slide presentation of one of the pliers. You should release the start button; then press an answer button as quickly and as accurately as you can. Then a slide will be presented so you can see if your answer was correct. Remember, you have only six seconds to press the answer button; and you should make the button press every time. If you don't know the correct answer, guess.

Please use one finger throughout the experiment. Which finger will you use? (RECORD HAND AND FINGER)

Are there any questions before we begin? (PAUSE--FOR ANY QUESTIONS, PARAPHRASE WHAT YOU'VE ALREADY SAID)

At first it will seem that the trial sequence happens quite rapidly and it will be easy to get out of sequence with it. You will know this has happened if there are long periods (7 seconds) with nothing on the screen followed by a slide of the pliers and the name. You can get back in sequence by pressing the "start" button when you see the pliers and the name and holding it down until a slide with just the pliers is presented. Once back in sequence you should make your response as fast and as accurately as you can and then <u>return</u> to the "start" button so you don't get out of sequence again.

OK, please put on the earphones and let's begin.

Group R_{MK}

Please have a seat in the chair. (POINT AND LET S GET SETTLED). In this experiment, your task is to learn to identify each of 8 different pliers by name.

Also, you will be asked to indicate whether you are "sure" or "not sure" of each answer that you give. A picture of a pair of pliers will appear on this screen. (POINT) First, you would press a button in this answer row to indicate the name of the pliers (POINT SWEEPINGLY), then you would press one of these two buttons to indicate whether you are "sure" or "not sure" of it. (POINT) You should try to press the buttons as quickly and as accurately as you can in the order that I said--first the answer button then the "sure" or "not sure" button. You should make an honest estimate of how sure you are of your answer, each time.

When we start the experiment, you will put on these earphones. (HOLD HEADSET UP) During the experiment, each trial will take place in the following order: First, you must hold this "start" button down. (POINT AND PRESS) You will hear a short tone through the earphones which will inform you that the pliers is about to be presented. After the tone, the pliers will appear on the screen and stay on for 6 seconds. During the 6 seconds, you have to release the start button, press the answer first; then press one of the "sure-not sure" buttons as fast and as accurately as you can. Shortly after you have pressed the buttons in the correct order, the picture will go off; and, then the same pliers will appear again on the screen together with the correct answer for 4 seconds so that you can see whether your answer was correct.

Shortly after, the brief tone will sound again to signal that the next pliers will be presented. You should make sure that you are pressing the "start" button when you hear the tone and that you hold it down until the slide is presented. If you don't do this, a second tone will be presented, then, if you don't press the "start" button immediately, a wrong answer will be recorded and you will be shown only the pliers and the correct answer after a 7 second period with nothing on the screen.

When the experiment first begins, the 8 pictures of the pliers along with their correct names will be presented one time for 5 seconds each so that you can study them. From then on, the pictures will appear alone and you will have to name them by pressing the proper answer buttons.

The 8 pictures of pliers will be repeated in a different order each time, until you have gone through

the sequence twice in a row without an error. The start of each sequence will always be indicated by a dot on the screen.

It will help you to tell the difference between the pliers if you notice that there are 2 different heads, 2 different shapes of handles and the handles may be plain or cushioned. (SHOW THREE PICTURES OF PLIERS ALONE FOR 45 SECONDS, POINTING OUT TO S THE CRITICAL CUES--THREE 8 x 10's THAT CONTAIN ALL FEATURES --RANDOMIZATION LISTED ON SCHEDULE)

Once again: the 8 pliers along with their correct names will be presented first and you should study them. Then your task begins with a dot on the screen and your finger on the "start" button. The brief tone sounds, followed by a slide presentation of one of the pliers. You should release the start button, press an answer and then one of the "sure-not sure" buttons, as quickly and as accurately as you can. Then a slide will be presented so you can see if your answer was correct. Remember, you have only 6 seconds to press the 2 buttons, and you should make the 2 button presses every time. If you don't know the correct answer, guess; and try to estimate accurately whether you are "sure" or "not sure" of each answer and this is very important!

Please use one finger throughout the experiment. Which finger will you use? (RECORD HAND & FINGER)

Are there any questions before we being? (PAUSE--FOR ANY QUESTION, PARAPHRASE WHAT YOU'VE ALREADY SAID)

At first it will seem that the trial sequence happens quite rapidly and it will be easy to get out of sequence with it. You will know this has happened if there are long periods (7 seconds) with nothing on the screen followed by a slide of the pliers and the name. You can get back in sequence by pressing the "start" button when you see the pliers and the name and holding it down until a slide with just the pliers is presented. Once back in sequence you should make your responses as fast and as accurately as you can and then return to the "start" button so you don't get out of sequence again.

OK, please put on the earphones and let's begin.

Group R_{MX}

Please have a seat in this chair. (POINT & LET S GET SETTLED) In this experiment, your task is to learn to identify each of 8 different pliers by name. A picture of a pair of pliers will appear on this screen. (POINT) First, you would press a button in this answer row to indicate the name of the pliers (POINT SWEEPINGLY), then you would press this button (POINT) to record your answer on our automatic recording system. You should try to press the buttons as quickly and as accurately as you can in the order that I said--first the answer button and then the "record" button.

When we start the experiment, you will put on these earphones. (HOLD HEADSET UP) During the experiment, each trial will take place in the following order: First, you must hold this "start" button down. (POINT & PRESS) You will hear a short tone through the earphones which will inform you that the pliers is about to be presented. After the tone, the pliers will appear on the screen and stay on for 6 seconds. During the 6 seconds, you have to release the start button, press the answer button first; then press the "record" button as fast and as accurately as you can. Shortly after you have pressed the buttons in the correct order, the picture will go off; and, then the same pliers will appear again on the screen, together with the correct answer for 4 seconds so that you can see whether your answer was correct.

Shortly after, the brief tone will sound again to signal that the next pliers will be presented. You should make sure that you are pressing the "start" button when you hear the tone and that you hold it down until the slide is presented. If you don't do this, a second tone will be presented, then if you don't press the "start" button immediately, a wrong answer will be recorded and you will be shown only the pliers and the correct answer after a 7 second period with nothing on the screen.

When the experiment first begins, the 8 pictures of the pliers, along with their correct names will be presented one time for 5 seconds each so that you can study them. From then on, the pictures will appear alone and you will have to name them by pressing the proper answer buttons.

The 8 pictures of pliers will be repeated in a different order each time, until you have gone through

the sequence twice in a row, without an error. The start of each sequence will always be indicated by a dot on the screen.

It will help you to tell the difference between the pliers if you notice that there are 2 different heads, 2 different shapes of handles and the handles may be plain or cushioned. (SHOW THREE PICTURES OF PLIERS FOR 45 SECONDS, POINTING OUT TO S THE CRITICAL CUES--THREE 8 x 10's THAT CONTAIN ALL FEATURES--RANDOMIZATION LISTED ON SCHEDULE)

Once again: the 8 pliers along with their correct names will be presented first and you should study them. Then your task begins with a dot on the screen and your finger on the "start" button. The brief tone sounds, followed by a slide presentation of one of the pliers. You should release the start button, press the answer button and then the "record" button as quickly and as accurately as you can. Then a slide will be presented so you can see if your answer was correct. Remember, you have only 6 seconds to press the 2 buttons and you should make the 2 button presses <u>every time</u>. If you don't know the correct answer, guess.

Please use one finger throughout the experiment. Which finger will you use? (RECORD HAND AND FINGER)

Are there any questions before we begin? (PAUSE--FOR ANY QUESTIONS, PARAPHRASE WHAT YOU'VE ALREADY SAID)

At first it will seem that the trial sequence happens quite rapidly and it will be easy to get out of sequence with it. You will know this has happened if there are long periods (7 seconds) with nothing on the screen followed by a slide of the pliers and the name. You can get back in sequence by pressing the "start" button when you see the pliers and the name and holding it down until a slide with just the pliers is presented. Once back in sequence you should make your responses as fast and as accurately as you can and then <u>return</u> to the "start" button so you don't get out of sequence again.

OK, please put on the earphones and let's begin.

Group R_{KM}

Please have a seat in this chair. (POINT AND LET S GET SETTLED) In this experiment, your task is to learn to identify each of 8 different pliers by name. Also, you will be asked to indicate whether you are "sure" or "not sure" of each answer that you give. A picture of a pair of pliers will appear on this screen. (POINT) First, you would press one of these two buttons to indicate whether you are "sure" or "not sure" (POINT) of the answer, then you press a button in this answer row to indicate the name of the pliers. (POINT SWEEPINGLY) You should try to press the buttons as quickly and as accurately as you can in the order that I said--first the "sure" or "not sure" button and then, the answer button. You should make an honest estimate of how sure you are of your answer <u>each time</u>.

When we start the experiment, you will put on these earphones. (HOLD HEADSET UP) During the experiment, each trial will take place in the following order: First, you <u>must</u> hold this "start" button down. (POINT & PRESS) You will hear a short tone through the earphones which will inform you that the pliers is about to be presented. After the tone, the pliers will appear on the screen and stay on for 6 seconds. During the 6 seconds, you have to release the start button, (11-20) press one of the "sure-not sure" buttons first; then press the answer button as fast and as accurately as you can. Shortly after you have pressed the buttons in the correct order, the picture will go off; and, then the same pliers will appear again on the screen together with the correct answer for 4 seconds so that you can see whether your answer was correct.

Shortly after, the brief tone will sound again to signal that the next pliers will be presented. You should make sure that you are pressing the "start" button when you hear the tone and that you hold it down until the slide is presented. If you don't do this, a second tone will be presented, <u>then</u> if you don't press the "start" button immediately, a wrong answer will be recorded and you will be shown only the pliers and the correct answer after a 7 second period with nothing on the screen.

When the experiment first begins, the 8 pictures of the pliers, along with their correct names will be presented one time for 5 seconds each so that you can study them. From then on, the pictures will appear alone and you will have to name them by pressing the proper answer buttons.

The 8 pictures of pliers will be repeated in a different order each time, until you have gone through the sequence twice in a row without an error. The start of each sequence will always be indicated by a dot on the screen.

It will help you to tell the difference between the pliers if you notice that there are 2 different heads, 2 different shapes of handles and the handles may be plain or cushioned. (SHOW THREE PICTURES OF PLIERS FOR 45 SECONDS, POINTING OUT TO S THE CRITICAL CUES--THREE 8 x 10's THAT CONTAIN ALL FEATURES--RANDOMIZATION LISTED ON SCHEDULE)

Once again: the 8 pliers along with their correct names will be presented first and you should study them. Then your task begins with a dot on the screen and your finger on the "start" button. The brief tone sounds, followed by a slide presentation of one of the pliers. You should release the start button, press one of the "sure-not sure" buttons and then an answer button as quickly and as accurately as you can. Then a slide will be presented so you can see if your answer was correct. Remember, you have only 6 seconds to press the 2 buttons; and you should make the 2 button presses every time. If you don't know the correct answer, guess; and try to estimate accurately whether you are "sure" or "not sure" of each answer and this is very important!

Please use one finger throughout the experiment. Which finger will you use? (RECORD HAND & FINGER)

Are there any questions before we begin? (PAUSE--FOR ANY QUESTIONS, PARAPHRASE WHAT YOU'VE ALREADY SAID)

At first it will seem that the trial sequence happens quite rapidly and it will be easy to get out of sequence with it. You will know this has happened if there are long periods (7 seconds) with nothing on the screen followed by a slide of the pliers and the name. You can get back in sequence by pressing the "start" button when you see the pliers and the name and holding it down until a slide with just the pliers is presented. Once back in sequence you should make your responses as fast and as accurately as you can and then return to the "start" button so you don't get out of sequence again.

OK, please put on the earphones and let's begin.

Group R_{XM}

Please have a seat in this chair. (POINT & LET S GET SETTLED) In this experiment, your task is to learn to identify each of 8 different pliers by name. A picture of a pair of pliers will appear on this screen. (POINT) First, you would press this button to set our automatic recording system to record your answer, then you would press a button in this answer row to indicate the name of the pliers. (POINT SWEEPINGLY) You should try to press the buttons as quickly and as accurately as you can in the order that I said--first the "record" button and then, the answer button.

When we start the experiment, you will put on these earphones. (HOLD HEADSET UP) During the experiment, each trial will take place in the following order: First, you must hold this "start" button down. (POINT & PRESS) You will hear a short tone through the earphones which will inform you that the pliers is about to be presented. After the tone, the pliers will appear on the screen and stay on for 6 seconds. During the 6 seconds, you have to release the start button, press the "record" button first; then the answer button as fast and as accurately as you can. Shortly after you have pressed the buttons in the correct order, the picture will go off; and, then the same pliers will appear again on the screen together with the correct answer for 4 seconds so that you can see whether your answer was correct.

Shortly after, the brief tone will sound again to signal that the next pliers will be presented. You should make sure that you are pressing the "start" button when you hear the tone and that you hold it down until the slide is presented. If you don't do this, a second tone will be presented, then if you don't press the "start" button immediately, a wrong answer will be recorded and you will be shown only the pliers and the correct answer after a 7 second period with nothing on the screen.

When the experiment first begins, the 8 pictures of the pliers, along with their correct names will be presented one time for 5 seconds each so that you can study them. From then on, the pictures will appear alone and you will have to name them by pressing the proper answer buttons.

The 8 pictures of pliers will be repeated in a different order each time, until you have gone through

the sequence twice in a row, without an error. The start of each sequence will always be indicated by a dot on the screen.

It will help you to tell the difference between the pliers if you notice that there are 2 different heads, 2 different shapes of handles and the handles may be plain or cushioned. (SHOW THREE PICTURES OF PLIERS FOR 45 SECONDS, POINTING OUT TO S THE CRITICAL CUES--THREE 8 x 10's THAT CONTAIN ALL FEATURES--RANDONIZATION LISTED ON SCHEDULE)

Once again: the 8 pliers along with their correct names will be presented first and you should study them. Then your task begins with a dot on the screen and your finger on the "start" button. The brief tone sounds, followed by a slide presentation of one of the pliers. You should release the start button, press the "record" button and then an aswer button as quickly and as accurately as you can. Then a slide will be presented so you can see if your answer was correct. Remember, you have only 6 seconds to press the 2 buttons; and you should make the 2 button presses every time. If you don't know the correct answer, guess.

Please use one finger throughout the experiment. Which finger will you use? (RECORD HAND AND FINGER)

Are there any questions before we begin? (PAUSE--FOR ANY QUESTIONS, PARAPHRASE WHAT YOU'VE ALREADY SAID)

At first it will seem that the trial sequence happens quite rapidly and it will be easy to get out of sequence with it. You will know this has happened if there are long periods (7 seconds) with nothing on the screen followed by a slide of the pliers and the name. You can get back in sequence by pressing the "start" button when you see the pliers and the name and holding it down until a slide with just the pliers is presented. Once back in sequence you should make your responses as fast and as accurately as you can and then return to the "start" button so you don't get out of sequence again.

OK, please put on the earphones and let's begin.