R E P O R T R E S U M E S

ED 020 654 56 EM 000 221

AN INVESTIGATION OF THE APPLICABILITY OF MOTION PICTURES TO EDUCATIONAL TESTING. FINAL REPORT.

BY- VAN HORN, CHARLES
ILLINOIS UNIV., URBANA

REPORT NUMBER BR-5-0713 PUB DATE 31 AUG 67
REPORT NUMBER NDEA-7A-665
GRANT OEG-7-23-1020-125
EDRS PRICE MF-$0.25 HC-$1.60 38P.

DESCRIPTORS- EDUCATIONAL TESTING, *FILMS, FACTOR ANALYSIS, EVALUATION, VISUAL STIMULI, EDUCATIONAL EXPERIMENTS, *TESTING, *COMMUNICATION (THOUGHT TRANSFER), VERBAL COMMUNICATION, SEX DIFFERENCES, AUDITORY VISUAL TESTS, PERSONALITY TESTS, VERBAL STIMULI, ECONOMIC STATUS, GUILDFORD ZIMMERMAN TEMPEPAMENT SURVEY

EIGHT ITEMS THE "GUILFORD-ZIMMERMAN TEMPERAMENT SURVEY" (E.G. , "YOU MAKE DECISIONS ON THE SPUR OF THE MOMENT") WERE DRAMATIZED AND FILMED IN TWO POLARIZED SITUATIONS, ONE DEPICTING A SPECIFIC SET OF BEHAVIOR TO SUPPORT AN AFFIRMATIVE RESPONSE TO THE TEST ITEM AND THE OTHER DEPICTING A NEGATIVE SPECIFIC. TWO ADDITIONAL VARIABLES INTENDED TO INFLUENCE RESPONSES TO THE PICTORIAL ITEMS WERE PRESENCE OF A MAN OR WOMAN AND AMOUNT OF MONEY IMPLIED IN THE DECISION SITUATION DEPICTED IN THE PICTURE. DUE TO COMPLICATIONS IN ADMINISTERING THE WRITTEN AND CINEMATIC FORMS OF THE TEST TO FRESHMEN, AND BECAUSE THE "UNDECIDED" RESPONSES WERE OMITTED, THE BASE FOR CORRELATIONS VARIED. HYPOTHESIS I, THAT "THE COMPLEXITY OF FACTORS IDENTIFIED FROM A SET OF TEST ITEMS INCREASES WHEN THEY ARE PRESENTED TO THE RESPONDENT IN THE FORM OF PICTURES INSTEAD OF WORD," WAS CONFIRMED, BUT THE COMPLEXITY WAS NOT EXPLAINED. HYPOTHESIS II, THAT "THE NATURE OF THE CHANGE IN FACTORIAL CONTENT OF A SET OF ITEMS REPRESENTED PICTORIALLY IS CONTROLLED BY THE EXPERIMENTER THROUGH HIS CHOICE OF THE PICTURES USED," HAD SLIGHT SUPPORT. SEX DIFFERENCES WAS THE MORE DEPENDABLE EXPERIMENTAL VARIABLE, THUS DISPELLING THE ASSUMPTION THAT INSTANCES OF THE EXPERIMENTAL VARIABLES WOULD BE EQUALLY ATTRACTIVE TO THE SUBJECTS. (RS)

FINAL REPORT
Project No. 665
Grant No. ~~711125~~
7-23-1020-125

AN INVESTIGATION OF THE APPLICABILITY OF MOTION PICTURES TO EDUCATIONAL TESTING

August 1967

U.S. DEPARTMENT OF
HEALTH, EDUCATION AND WELFARE

Office of Education
Bureau of Research

An Investigation Of The Applicability of Motion Pictures To Educational Testing

Project No. 665
Grant No. ~~711125~~ 7-23-1020-125

Charles Van Horn

August 31, 1967

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

University of Illinois

Urbana, Illinois

Contents

Tables

An Investigation of the Applicability of Motion Pictures to Educational Testing

Charles Van Horn

I. Introduction

Time and repetition have reduced to a truism the idea that effective teaching utilizes two avenues of communication. They have not invalidated the idea nor diminished the consequences of flouting it, nor even simplified the achievement of the aim that it expresses. The ability to say and do things that will modify his pupils' behavior is still a part of the measure of a teacher's effectiveness, and his ability to observe behavior and determine whether the modifications he detects are the ones he intended are another measure of it.

Instruction influences evaluation and evaluation influences instruction because the two take place sequentially, and the interaction between them is so close that neither is likely ever to be more effective than the other. So long as pupil and teacher are in close and frequent contact with one another the interdependence of teaching and evaluating never becomes a problem. A system which undertakes to educate larger numbers of pupils is necessarily a more complex one, and probably a more impersonal one because the opportunity for interpersonal contact is lost.

When he is deprived of the opportunity for extended observation of his pupils, the teacher is forced to rely on a smaller number of pre-arranged and systematically organized observations. Testing -- large scale testing -- has entered the educational process.

Teaching and testing both are complex activities. Both are highly specialized and unfortunately, seldom seem to be matters of primary concern to the same persons so that a community of teachers exists along with a community of testers. Ideally, any procedure or strategem utilized by practitioners in either field should influence the activities of the other, but practice often falls short of this ideal.

One instance in which coordinate growth in the two aspects of education seems to have lagged is in the use of motion pictures. As instructional devices, motion pictures are anything but recent innovations; their place in teaching is so widely accepted and their use so freely adopted that documenting their origin and growth is not attempted here.

The concomitant development of tests which use motion pictures to systematize the observation of pupils' behavior has not occurred. In the face of widespread use of motion pictures for

instruction, formal measurement methods continue to rely almost entirely on verbal -- specifically, paper-pencil -- instruments.

An anomolous situation exists in which educators seem to be saying that some ideas are of such a nature that words are not adequate for their expression, but that those same inadequate words can be used to appraise the pupil's understanding of the nonverbalizable ideas.

An interesting question remains unasked about the nature of measurement when it is implemented by written words and when it is implemented by pictures. Tremendous differences exist between words and pictures in the way they function as linguistic devices; differences so great that, at this point, it cannot be assumed that the same question can be asked in both ways.

Language is man's great achievement and it owes its power to two characteristics: it is largely arbitrary and it is completely abstract.

Arbitrariness is revealed by the fact that words bear no resemblance to the things or ideas to which they refer. A picture, by contrast, conveys its meaning largely by virtue of its resemblance to its referent.

Arbitrariness also appears in the existence of more than one combination of words which may be used to refer to a given thing ("thing" refers here, not only to physical objects, but to conditions, events, activities, ideas, etc.). A speaker always makes his choice from a wide range of possible sequences and the nature of the idea seldom dictates a unique form of expression. It is possible to provide an alternative verbalization which will convey almost, but never exactly, the same meaning.

The lack of complete correspondence exists because of the second important characteristics of words, their abstractness. A word does not typically refer to a thing, but to some abstracted property which is common to all members of some class of things to which the selected one belongs. Any object or idea belongs to a number of classes and could legitimately be referred to by any of the properties of any of those classes.

Unequivocal designation is achieved only by adding other words to name other classes to which this thing belongs, or to exclude unwanted classes which share an unwanted property, or to designate intersections of this class with others. Eventually, if both speaker and listener are patient and attentive, the original abstraction can be reduced to the single object or event or condition originally intended.

Compare this process, cumbersome in the telling, but expertly

executed by every individual every day throughout most of his life, with a picture yields a higher degree of specificity with a glance.

The individual who attempts to communicate by the use of pictures confronts the same problem in a different context; instead of delimiting an over-inclusive idea by adding more words, he expands on an over-specific one by adding more pictures.

This is not intended to imply that communication by means of pictures is always desirable or that it is inherently superior to verbal expression. The intent is to demonstrate that each mode posesses characteristics which make it suitable for conveying certain kinds of ideas. If the speaker's concern is with abstract ideas, the nonabstract nature of pictures would be a liability and, if his concern is with specific behaviors, the nonspecific nature of words becomes a liability.

Before confronting the question of the possibility of asking the same question with both kinds of expression, the distinction between the nature of words and the nature of pictures needs to be examined for its implications to the construction of tests.

The testing procedure is regarded as resting on a common understanding, i.e., accurate communication, between two individuals, the test-maker and the test-taker. Unless the latter forms an accurate impression of the materials provided by the former, his responses will not have the meaning that is imputed to them and the test will not have measured that which was intended.

Imagine a continuum of communicative efforts, representing all possible degrees of accuracy. One end might be represented by a problem in a mathematics examination which is not likely to be misunderstood because it is phrased in a highly specialized language designed to emphasize singularity of meaning and rigidity of syntax. The other end might be a projective test of the inkblot variety in which the stimulus is amorphous and is purposely expressed in a form which has neither vocabulary nor syntax in order to assure maximum equivocality. Somewhere between these two, there lies the case of the typical examination in which the question is phrased in unspecialized language, partially structured but subject to the peculiarities of interpretation already described.

The exact location of that "typical" test along the continuum is yet to be specified. If it is near the precise end, then the answer to the question about feasibility of interchanging the two ways of asking a question should be an affirmative one, if it is near the unstructured end, the answer should be negative.

The problem becomes that of devising an experiment which will permit differences between the two kinds of expression to assert themselves if they exist, without limiting the possibility of demonstrating a correspondence between them. The apparent conflict between these two objectives is resolved by resort to the famous dictum of psychological experiments, that if one wants to understand something, he must watch it change.

The procedure must be that of asking the same question in several ways which differ from one another according to the amount and kind of detail they provide. If the meaning attributed to a statement by a respondent is independent of its form, then all of the questions should be answered in the same way; if meaning is influenced by form, then responses the different versions should differ; and if the iconic nature of pictures can be exploited by the experimenter, the differences should be predictable.

A temperament measure provides a suitable vehicle for investigation of the differences between words and pictures. First of all, it provides a source of questions which are grouped according to content, so that each demonstrates internal consistency. This facilitates inter-item comparisons in analysis. Second, it consists of brief descriptions of many kinds of situations to provide the respondent with opportunities to describe the way he would act or feel. Third, the descriptions are phrased in general terms because generalized reaction tendencies are being sought. This third condition is of primary importance because the distinction between general and specific utterances is crucial to the experiment.

Examination of the vocabulary used in such questionnaires shows that many pronouns are used because the questions are addressed to specific persons. They make liberal use of verbs because they deal with behavior; many nouns are found because the behavior must involve some agent acting in some context. Adverbs and adjectives are used less frequently because they would increase the specificity of the description and specific questions would draw specific responses which would limit the allowable inference to the range of situations described.

The third circumstance, absence of specificity in expression, is an important one because it provides the experimenter with more opportunities to alter the question in the process of conversion into a picture. The nature of pictorial expression is such that most of the nouns, many of the verbs, and almost every adjective and adverb can be manipulated by the experimenter. Pictorial details, not present in the verbal statement, may be added at his discretion to make specific that which was formerly general.

Appeals to differing, and possibly conflicting, reaction tendencies can be introduced into a picture in order to provide a

basis for comparison with the parent verbalized general statement.

The success of the experiment depends on the skill with which the experimenter devises pictures which, although they represent a situation which is describable by the verbalized question, depict details not present in the verbal statement which appeal to reaction tendencies more consistently organized than the one the original statement was intended to measure.

A hypothetical example can be used to demonstrate the way in which the calculated introduction of specific elements into a general statement might influence responses. Consider the verbally stated question, "Do you habitually interrupt others in conversation?", and suppose that it could be answered truthfuly and without difficulty by most persons.

Now suppose that a pictorial sequence is prepared which will constitute an instance of the kind of behavior described, and that the same persons are asked to respond to the question asked in the new way. In the process of preparing this picture, the necessity arises for depicting some specific "other" whose conversation is to be interrupted.

Suppose further that the experimenter displays to a respondent who answered the verbal question affirmatively, a pictorial sequence which depicts him, as a student, in conversation with a professor. An extraneous element has been added to the question; it was not present in the generalized (verbal) statement, but neither was it excluded by that statement. The response to this form of the question would be compounded of the respondent's tendency to specify that interruption of others characterizes his behavior, of the experimenter's skill in depicting the difference between the academic status of the two individuals, and the respondent's tendency to defer to professors.

If the experimenter had asked the pictorial question in a different way, for example, had he shown a group of students in casual conversation, the effect of the added element might have been different. The two pictorial forms have in common the inclusion of an element which was not overtly mentioned in the verbal statement, and they differ in the specific instance of that element.

Neither of the picture questions could be the basis for a valid inference concerning the trait dealt with by the verbal form because they can be generalized only to cases in which the interrupted conversation involves a person whose status, relative to that of the respondent, has been specified. Any difference between the responses to those two pictorial questions could, how-

ever, become the basis for an inference regarding that respondent's ability to percieve and tendency to react to status differences in social behavior.

An extraneous element, invented by the experimenter, has been introduced into the question. The experimenter who systematically introduces such extraneous elements, differing in nature and intensity, over a series of questions may be able to demonstrate systematic differences between responses to verbal and pictorial questions.

Considered as an experiment, the independent variable in such a procedure would be the prominence accorded to the extraneous elements which are introduced into pairs of questions. The dependent variable would be the factorial structure of the set of questions presented. The mechanism is the difference between the specificity present in verbal and pictorial statements. The limitation is the skill of the experimenter in devising and presenting credible instances of the behaviors he selects. The permitted inference is the nature of the difference between verbal and pictorial expression and, hence, the feasibility of substituting one for the other in educational testing.

From the consideration of the difference between words and pictures and of the proposed manipulation of the difference, these hypotheses may be adduced:

> Hypothesis 1. The factorial complexity of a set of items taken from a temperament survey increases when they are presented to the respondent in the form of motion pictures instead of words.
>
> Hypothesis 2. The nature of the change in factorial content of a set of test items, when they are pictorially represented, is controlled by the experimenter through his choice of the content of the pictures.

The phrasing of these hypothesis , in preference to the conventional null form, is that of a positive statement. Since no test of a null hypothesis is feasible in a factor analytic experiment, no advantage is attached to its use and its phrasing might obscure the intent of the experiment.

History and Precedent

The possibility that motion pictures might have unexpected advantages in testing is not a new one, but it seems to have received relatively little attention, at least in proportion to the number of words that have been devoted to it. Gibson (1947)

seems to imply a negative answer to the question which this experiment deals in his statement:

> ". . .the motion picture makes available to the test designer not only a special method of measuring known factors of ability, but also gives him access to new and unnamed functions not accessible to conventional methods of test construction."

Carpenter (1954) is more specific, but not more optimistic. He observes:

> ". . .the advantages of film testing may appear in the assessment of personality characteristics where the observation of responses to cur ently used instruments is confused by constantly changing stimuli or by the use of stimuli which must be ambiguous to be both complex and constant."

Despite these expressions of confidence by men prominent in their fields, so little research of the kinds they suggest has been reported that the experiment described here is almost without precedent. The absence of experiments similar to this one does not indicate that psychologists have overlooked motion pictures or that they have ignored their existence, but that their attention has been directed toward characteristics of films other than the ones with which this experiment is concerned. This experiment seeks to exploit differences between verbal and pictorial representation of complex ideas, while most previous ones have chosen to accept, rather than to manipulate, the iconic nature of pictures.

Some idea of the extent of psychologist's concern with motion pictures and of the direction their concern has taken may be obtained by examining references to them in the Psychological Abstracts. During the fifteen year period between 1947 and 1962, the Abstracts carried 409 entries under the headings of Cinema, Film, and Motion Pictures; these were distributed in this way:

- 234 reviews of films pertaining to educational or psychological topics.
- 80 discussions of films in which no experiment was performed and no data gathered.
- 71 reports of experiments in which a film was used for teaching or training.
- 10 reports of experiments dealing with audience reactions or behavior.
- 7 reports of experiments in which a film was used for a purpose other than entertainment, education, or persuasion.

7 references not classifiable under one of these categories.

Of the seven which depart from the conventional uses of motion pictures, the work of Gibson (1947) and of Carpenter et al (1954) concerns itself with measurement of perceptual skills and subject-matter knowledge.

Not enough is known about the use of Lhotsky (1953,1955) to permit an evaluation of its pertinence, but he seems to have used motion picture sequences as stimulus material for projective measurements. Spiegelman (1956) followed a similar approach, but asked viewers to make judgements about the personality of the film-maker. These experiments have little direct bearing on this one, but they represent isolated instances of the use of motion pictures inan experimental context.

Johnson and Vogtmann (1955) encountered an interesting circumstance when they used a twenty minute sequence from a commercially produced entertainment film as part of a final examination in a class in introductory psychology. A series of multiple choice questions were asked about the interpersonal relationships demonstrated, motivations of the characters, etc. Students' reactions were favorable and internal consistency reliability was high, but a subsequent administration demonstrated that entering students performed only slightly less well than those who had completed the course. The authors conclude that either the entering students were very skillful at analyzing personal relationships, or that the course has little in common with the test.

They also suggest, however, that improvement might result if teaching were done in the same way as testing -- a statement that seems particularly relevant in this context. In any event, the suggestion is present that their test called on skills or knowledges not included in their instruction and, hence, that the use of films in educational testing introduces some factor or element not included in conventional verbal examinations.

Stoller and Geertsma (1958) produced an interesting test in which they used motion pictures to evaluate the clinical skills of medical students in psychiatry. The students observed a thirty minute film of a patient and then answered 100 questions regarding his disorder. They made no attempt to manipulate any variables associated with their pictures, but they offer the interesting observation that information obtained in this way about the performance of the students provides a basis for revising methods of instruction.

Neither of these experiments attempted to investigate the effects of altering the content of the pictures, and both used

long films which would make it difficult to demonstrate the effect of any such changes. Both are relevant here, however, because they show direct relationships between the nature of instruction offered and the nature of the test used to evaluate that instruction.

An experiment reported by McIntyre (1954) is directly relevant to this one because he used short motion picture sequences which were based on cards from the Thematic Apperception Test and compared pairs of responses to them. He prepared two films based on each of five cards, showed them to respondents who resembled the actors either in age or sex, and asked questions about the action they had seen.

He concluded that the realistic nature (specificity) of the films may be in opposition to the ambiguity necessary for projection, and the action depicted (extraneous elements added) may further structure the situation to reduce ambiguity. These conclusions contain two implications for the experiment considered here, and both of them point toward the expectation of little projection by the viewers in this experiment.

The most direct precedent to this experiment is described in an unpublished report of a test developed by Rose et al (1953) at the University of Southern California. Twenty-two very short film sequences were prepared, each presenting a separate conflict situation and two alternative solutions to it. The subjects, ROTC cadets, were first shown a short film, one of four prepared for this purpose, in which the behavior of one individual was observed for approximately eight minutes. They were then required to specify, for each of the twenty-two problem situations, the kind of solution that would be adopted by the individual they had observed. The question of the respondent's own behavior was never raised.

The results indicated clearly that a stable conception of the observed individual was formed by the subjects (about 1100 cadets divided into four groups) in each of the four versions shown, but that some were percieved more uniformly than others. The authors postulate that their test measures a trait of social perception in a way that avoids the problems of falsification and inability to verbalize responses.

The similarity between that experiment and the one reported here is close. They have in common the use of filmed social situations revolving about some conflict in order to produce a controlled amount of ambiguity, and the fact that the conflicts in both were suggested by statements from the Guilford-Zimmerman Temperament Survey. The two experiments differ in three important ways: (1) Rose's items were concerned only with the Ascendancy dimension from the Survey, while this experiment deals with two

dimensions (Ascendancy and Restraint); (2) the distracting elements in Rose's experiment were not systematically varied according to intensity; (3) responses were made in terms of prediction of another person's behavior.

Most of the experiments cited here are of interest only because they represent instances of the application of motion pictures to specific measurement problems which are identified with specific instructional sequences. Those of Rose and McIntyre of special interest because they utilize series of film sequences made for experimental purposes and because they embody experimental variables introduced in such a way that their effects can be isolated in analysis.

Of the two, McIntyre's experiment is the more carefully constructed; Rose's experiment represents a larger departure from conventional testing procedure and is more important for the suggestions it yields than for the quantitative data it produces. This experiment has been devised to extend the idea contained in Rose's experiment and to test the hypotheses implicit in it.

II. Method

This experiment can be regarded as consisting of five steps:

(1) Selection of the extraneous elements to be introduced into the experiment. Those selected were amounts of money and the presence of girls; this selection was based on the intent to use male undergraduate college students as subjects.

(2) Identification of the verbal statements which would become the basis for the pictorial sequences. Those used were four items each from the Restraint and Ascendancy scales of the Guilford-Zimmerman Temperament Survey.

(3) Preparation of a pair of motion picture sequences based on each verbal item in such a way that one member of each pair depicted the presence of an extraneous variable, and the other its absence.

(4) Administration of both the verbal and the pictorial versions to a group of subjects.

(5) Comparison of the factorial structures of the two kinds of items with emphasis on the presence of effects attributable to the extraneous elements.

Selection of Experimental Variables

Given the intention to make motion pictures and to record responses of undergraduate men to those pictures, the selection of sex and money as experimental variables was not difficult. Both of these can be represented readily in films, and these respondents are assumed to be alert to the implications of the presence of either.

Identification of Parent Verbal Statements

The selection of The Temperament Survey as a source from which to draw verbal statements as a basis for experimental comparisons was based on several considerations.

The two scales selected demonstrate high internal consistency reliability (at least .8) and low intercorrelation (-.08). Their use increases the likelihood that two sets of questions can be found which will provide opportunity for the effects of the experimental variable to appear in two contexts, and which will provide an identifiable factorial structure.

Three steps were involved in the selection of the items to be used. In the first, a factor analysis of items from the two scales was conducted, using data from another experiment. This analysis helped to identify the items that would yield a satisfactory structure and was used as a guide in further selection.

The second step was the selection, from this set, of those items which could be converted into pictures. To qualify, the item must refer to some overt activity which can be photographed, preferably briefly and directly, without conversation or verbal description. After discarding those items which refer to the thoughts or intentions of the respondent, further selection was made by examining the nouns and adjectives in the statement, since the manipulation of the experimental variables is best accomplished by altering those ideas which name or describe.

Having identified the items for which visual counterparts might be prepared, the third step was made. This condition limits the choice to those items from which filmed counterparts may be made in pairs in such a way that one member displays the experimental variable more prominently than the other. Ideally, the manipulated variable should be absent, or nearly so, from one sequence, and very prominent in the other, but this ideal could not be achieved.

This was the most difficult part of the experiment and, it should be pointed out, represents a complication which is not inherent in film testing but is imposed by the nature of this experiment. The difficulty lay in the absence of an experimental precedent or a theoretical rationale which would estimate the strength of the distracting element which would have just the desired effect on the subjects' perception of the scene.

Since the expense of preparing these items prohibits the use of pretest data as a basis for discarding unsatisfactory ones, it becomes necessary to make two assumptions regarding these items without adequate assurance that they are being met. First, it is assumed that all of the printed items selected are of equal strength with respect to the appeal they make to the basic trait. Second, it is assumed that within each pair of pictorial items, the respondents' perceptions of the difference in attractiveness of the experimental variable match the experimenter's judgement.

Preparation of the Motion Picture Sequences

The third step in the experiment, producing the film sequences described below, took place at the University of Illinois in the spring of 1961. They were prepared with the aid of a grant from the Small Grant Program of Title VII of the National Defense

Education Act of 1958. Production was accomplished by Mr. Stanley Follis, at that time a member of the staff of the film unit of the University of Illinois High School. He was assisted by the experimenter and two students, both members of a film-production class. The actors were students recruited for this purpose.

The items chosen from the Temperament Survey and a description of the content of the film items associated with each of them is shown below.

Restraint items in which money is represented:

Survey item 67: "You make decisions on the spur of the moment."

Film item 67-1: Two men are being served in a coffee shop. One indicates his choice from among a large variety of doughnuts displayed as soon as he sits down. The other hesitates over a choice and, after being served, changes his mind and asks for a different doughnut.

Film Item 67-2: Two men are seated in a restaurant which appears to be an expensive one. One gives his order as soon as he looks at a menu. The other hesitates, gives an order, reexamines the menu and changes his order.

Survey item 147: "You are fond of betting on horse races and games whether you can afford it or not."

Film item 147-1: Three students are shown. One of them who is playing with some change in his hand speaks to another who shakes his head. He speaks (again without sound-track) to the third who takes coins from his pocket and they begin to match coins.

Film item 147-2: A close-up shows a pair of hands as they spread five cards; the cards are three sixes and two eights. Dissolve to a hand lingering over a small pile of chips and finally putting forth a few of them in a hesitant way; dissolve again to a hand which takes most of the chips from the same size pile and thrusts them forward confidently.

Items from the Restraint scale differing according to sex

Survey item 17: "You like the parties you attend to be lively."

Film item 17-1: The first scene shows two men in quiet conversation in a setting which suggests a private home, over a background of music and muted conversation. Cut to a scene which

shows three men laughing boisterously and talking loudly in a scene which suggests a tavern.

Film item 17-2: One scene shows a man and a woman talking quietly in a setting similar to that in the first part of 17-1; music is in the background and the scene fades as they rise to dance. Fade in to a group of people, dressed as if for a party, and engaged in a lively activity (charades).

Survey item 27: "You like work that requires considerable attention to detail."

Film item 27-1: In a book store, a man greets a girl as she enters; they engage in a conversation (not recorded) as they walk to a shelf where she points to a row of books. He takes one down, hands it to her for examination, and they walk together to a cash register, where another man sits. The first man turns away to greet another girl, the scene remains with the second who, paying only casual attention to the girl, engages in detailed computation before accepting the money and recording the transaction.

Film item 27-2: A photographer in a studio setting is working with an attractive model. He makes painstaking adjustments and pays little attention to the girl. Cut to a picture of a photographer watching a baseball game. He makes several pictures in quick succession without showing any concern for camera or technical details of its operation.

Ascendency items in which money is represented:

Survey item 8: "When you find that something you have bought is defective, you hesitate to demand an exchange or refund."

Film item 8-1: At a football stadium, a man buys a sack of peanuts from a vendor and finds, when the first one is opened, that it is not edible. He is shown, first returning to the vendor and asking for the return of his dime, then the scene dissolves to the alternative in which he throws the entire sack into a trash can and walks on.

Film item 8-2: A man is shown buying a record in a music store. The register indicates a price of $4.98. Dissolve to a close-up of the record being placed on a turntable and accompanying sound indicating that it is cracked. The alternatives shown are of a record being thrown into a wastebasket, and that of the same man securing a refund from the store.

Survey item 48: "You find it difficult to ask people for money or other donations, even for a cause in which you are interested."

Film item 48-1: Two men are collecting contributions from people as they leave a building. One holds a collection can which he shakes animatedly while speaking to the person he is facing, then turn immediately to another as soon as the first one places coins in the can. The other holds forth a similar can in a dispirited way while several persons walk past him, while he makes no effort to attract their attention or to stop them.

Film item 48-2: A man seated behind a desk in an office tells the two men facing him, "Our hospital fund is not going well enough. We've got to go after some big contributions if we're going to make our quota. Have you talked to Mr. Van Horn yet? He should be good for five thousand." The man addressed replies, "No, I don't think there is any use in asking him; he's hard to talk to and he's never given anything yet." The third man rises, remonstrates with the second and indicates his intention of confronting the recalcitrant donor. (Note that recorded conversation is used in this sequence, as in 88-2, to convey the intent of the actors).

Ascendency items differing according to sex:

Survey item 88: "When you are attracted to a person you have not met, you make an active attempt to get acquainted even though it may be quite difficult."

Film item 88-1: A man entering a classroom notices that two empty chairs remain; one is near the aisle between two men, the other is at the fartherest end of a row next to a girl who smiles and indicates the empty seat next to her. As he starts to crowd his way toward that seat, the other students (male) indicate displeasure and call his attention to the nearer empty seat. In one alternative he yields to this pressure and sits in the near seat between two men; in the other he continues, despite their protests, to the other end of the row and the seat near the girl.

Film item 88-2: A lecturer is shown and his closing remarks heard as he addresses a group of students. Two of the students whisper to one another, indicating their interest in what he has said. As the lecture ends, the group crowds around the lecturer, barring the two from access to him. One of the pair elbows his way through the crowd until he confronts the lecturer, the other turns away dejectedly.

Survey item 133: "You find it somewhat difficult to say 'No' to a salesman who tries to sell you something you do not really want."

Film item 133-1: Two men, carrying ice-cream cones, walk out of a store and are confronted by a magazine salesman who blocks their path. One walks around him and leaves the scene; the other makes

several unsuccessful attempts to get away and finally laboriously transfers both of his ice-cream cones to the other hand to reach for a coin which he hands to the salesman.

Film item 133-2: An attractive girl confronts two men on the street, offering crepe paper flowers for sale. Both men decline, the girl persists, one man walks on, the other offers more objections, but eventually buys a flower.

Having been photographed and edited to the shortest length deemed feasible, the items were arranged, by a random process, into the following order:

1. 48-1	5. 88-1	9. 17-1	13. 27-1
2.147-1	6.147-2	10. 67-2	14.133-2
3. 67-1	7. 48-2	11. 8-1	15. 17-2
4. 88-2	8. 8-2	12.133-1	16. 27-2

It will be noticed that each film sequence either depicts two people who behave in different ways, or one person who is shown resolving his conflict in each of two ways. A still photograph was made for each item depicting the characters in such a way that their important characteristic was demonstrated. These pictures were labeled "A" and "B" on the face of the photograph and approximately seven feet (fifteen seconds) of motion picture film of that still picture were added to each film sequence at the close of the action so that the courses of action were made explicit and a mode of response provided. At the end of each of these pictures, a seven second length of clear leader was inserted into the roll to provide time and illumination for marking answer sheets.

The following oral instructions preceded the showing of the film:

> "This film will show some people doing some very ordinary and routine things. Most of them are the kind of things you have probably done yourself at some time. In the film, every situation has two possible outcomes -- two ways in which it could have been done. As you see each sequence imagine that you are the person involved in it and decide which way you would behave. You will see both of the outcomes, so you will need only to decide which person did the thing you yourself would have done in that situation. After each scene is shown, the people will be identified on the screen by the letters A and B, then you can mark your answer sheet with an "A" or a "B" to show how you would have behaved."

(Show Item 48-1)

"If you were collecting contributions in this way, would you have behaved more like the man on the left or the one on the right? If you think you would have behaved more like the man on the left, mark an "A" in the space numbered "1" on your answer sheet; if you think you would have behaved more like the one on the right, mark a "B" in that space."

For the remaining fifteen items, the sound track posed the relevant question, i.e., "What would you do?" or "Which party would you rather attend?", etc., while the still picture associated with that item was on the screen. When the screen was illuminated by the clear leader, the further oral instruction was read "Mark item (insert item number) A or B.".

Administration Of The Two Forms Of the Test

The first data gathered for the experiment was the administration of the printed Temperament Survey to the entire freshman class of the Army ROTC Unit at the University of Illinois on May 1, 1961. The session in which this form was administered was conducted by a different person who conducted the film session so that no connection between the two incidents might be made by the subjects.

At the time the first data was taken, it was intended that the film items would be shown later in the month. Several circumstances, all of them external to the experiment, intervened and it was necessary to postpone the film session for nearly a year, until March 21, 1962. By that time, transfers and the attrition which normally accompanies the change from freshman to sophomore status had reduced the number of available subjects for whom Survey scores were recorded from 400 to 192.

For this reason, the correlations between printed items are based on approximately twice as many cases as those between film items or those between film and printed items.

Comparison of Factor Structures

The final step in the experiment was the intercorrelation and factorial analysis of the responses obtained in the two testing sessions. The results of these analyses and their comparisons and interpretations are shown in the chapters that follow.

III. Results

The first step in the analysis was the intercorrelation and factor analysis of the responses made to the printed form of the Temperament Survey. Item counts were compiled manually from the answer sheets, omitting from each comparison, all of the cases in which either item had been answered with a middle category (Undecided) response. Because of these omissions, the number of cases on which the correlations are based varies slightly (from 364 to 400).

The correlations, shown in Table I, were computed by a cosine-pi approximation to a tetrachoric coefficient from four-fold tables based on agreement, or lack of it, with the keyed responses. Had the responses been correlated according to "Agree" and "Disagree" categories, all of the correlations involving items 27 and 88 would have been reversed in sign.

Table I

Correlations between eight selected items
from the Guilford-Zimmerman Temperament Survey
(Decimal points are omitted)

Item	17	27	67	147	8	48	88	133	p
17		-139	466	303	-127	-208	-399	071	112
27	-139		102	100	140	019	120	149	515
67	466	102		372	160	-042	-169	-028	531
147	303	100	372		-063	-095	-079	-023	792
8	-127	140	160	-063		407	342	434	546
48	-208	019	-042	-095	407		338	482	453
88	-399	120	-169	-079	342	338		334	687
133	071	149	-028	-023	434	482	334		716

The column headed "p" in Table I refers to the proportion of keyed responses associated with each of the items in this sample.

The four centroid factors shown in Table II were extracted from this matrix and rotated manually to the structure shown in Table III.

When data had been obtained from the film testing session, the eight relevant responses from the appropriate Survey answer sheets were recorded on punched cards along with responses to the sixteen filmed items. Cosine-pi approximations of the correlations were computed manually because the only available machine program would have made it necessary to remove every case in which even one "Undecided" response appeared. The extra labor of manual

computation was considered preferable to the sacrifice of data that would have been discarded under this plan. The correlations obtained in this way are reported in Table IV; in interpreting these correlations it should be borne in mind that, unlike those shown above, these are computed directly from responses, without concern for the direction of keyed response. As a consequence, the signs assume importance in determining whether an item has functioned as anticipated, and the possibility of achieving a positive manifold in the analysis is abandoned.

Table II
Unrotated Factor Loadings

Item	A	B	C	D	h^2Est.	h^2Obt.
17	-650	505	-405	-072	848	847
27	145	172	277	217	174	174
67	-372	577	235	-198	568	566
147	-319	398	194	239	354	355
8	526	457	126	-328	606	609
48	532	264	-133	-109	390	382
88	581	108	143	089	378	378
133	595	535	-408	198	834	846
% V.	46.6	32.6	13.3	7.5		

Table III
Rotated Factor Loadings

Item	A	B	C	D	h^2
17	-042	510	-758	-072	842
27	162	174	265	217	174
67	074	713	-087	-198	561
147	-019	539	-074	239	353
8	656	103	218	-328	610
48	583	-157	071	-109	382
88	456	-164	364	089	375
133	867	-128	-197	198	846

The eight centroid factors shown in Table V were extracted from this correlation matrix. In the original extraction, the largest correlation in each row was used as an estimate of the communality for that variable; these results were iterated to stability with the exception that unit communalities were used in the second iteration for variables 17 and 8-2, despite the fact that communalities greater than one had been obtained. These variables were rotated with greater-than-unity communalities and, while this circumstance does not invalidate the experiment,

Table IV

Correlations Between Eight Printed Items
and Sixteen Filmed Items

Item	17	17-1	17-2	27	27-1	27-2	67	67-1	67-2	147	147-1	147-2
17		651	220	-068	-029	-205	405	114	-121	-006	126	-238
17-1	651		-269	214	-249	-124	-231	216	163	-217	-247	129
17-2	220	-269		-114	-167	155	154	177	-178	149	-002	-166
27	-068	214	-114		-227	-032	-091	-084	-100	154	-023	-006
27-1	-029	-249	-167	-227		-147	363	026	-416	134	056	015
27-2	-205	-124	155	-032	-147		-061	239	-165	-148	168	-063
67	405	-231	154	-091	363	-061		052	-051	358	119	-191
67-1	114	216	177	-084	026	239	052		505	-117	006	077
67-2	-121	163	-178	-100	-416	-165	-051	505		-128	-149	168
147	-006	-217	149	154	134	-148	358	-117	-128		266	-285
147-1	126	-247	-002	-023	056	168	119	006	-149	266		-052
147-2	-238	129	-166	-006	015	-063	-191	077	168	-285	-052	
8	150	-174	060	-227	155	-134	231	157	182	119	030	041
8-1	-139	-041	-217	227	106	-125	056	008	-286	-079	-107	140
8-2	-300	119	301	-114	317	-185	391	415	273	613	-184	-163
48	079	-049	-269	-066	-294	112	-129	-167	026	-105	-183	-026
48-1	-250	-205	118	-122	185	071	124	145	273	-087	108	-016
48-2	-139	173	-027	-165	-105	204	-011	255	062	-035	180	106
88	565	-025	107	078	237	175	193	170	017	079	171	-256
88-1	-474	087	-215	-029	-396	059	-287	-122	010	025	-295	070
88-2	093	-212	-055	-039	494	-016	049	005	-136	139	-037	-028
133	083	-089	-133	-179	021	341	016	278	223	154	017	-041
133-1	-385	-037	-037	024	225	012	067	137	222	-372	-138	104
133-2	-056	-027	127	-012	075	-030	-017	141	100	016	349	075

Table IV

Correlations Between Eight Printed Items
and Sixteen Filmed Items

	8	8-1	8-2	48	48-1	48-2	88	88-1	88-2	133	133-1	133-2
17	150	-139	-300	079	-250	-139	565	-474	093	083	-385	-056
17-1	-174	-041	119	-049	-205	173	-025	087	-212	-089	-037	-027
17-2	060	-217	301	-269	118	-027	107	-215	-055	-133	-037	127
27	-227	227	-114	-066	-122	-165	078	-029	-039	-179	024	-012
27-1	155	106	317	-294	185	-105	237	-396	494	021	225	075
27-2	-134	-125	-185	112	071	204	175	059	-016	341	012	-030
67	231	056	391	-129	124	-011	193	-287	049	016	067	-017
67-1	157	008	415	-167	145	255	170	-122	005	278	137	141
67-2	182	-286	273	026	273	062	017	010	-136	223	222	100
147	119	-079	613	-105	-087	-035	079	025	139	154	-372	016
147-1	030	-017	-184	-183	108	180	171	-295	-037	017	-138	349
147-2	041	140	-163	-026	-016	106	-256	070	-028	-041	104	075
8		-392	481	350	-032	454	-356	000	-148	437	-139	216
8-1	-392		-169	-172	072	-213	163	087	282	-201	239	-061
8-2	481	-169		204	100	278	-234	295	208	059	-009	169
48	350	-172	204		-279	315	-297	078	-110	306	-165	-008
48-1	-032	072	100	-279		-188	-068	055	184	-115	048	-037
48-2	454	-213	278	315	-188		-335	124	-385	401	-080	-004
88	-356	163	-234	-297	-068	-335		-210	342	-049	073	101
88-1	000	087	295	078	055	124	-210		-021	051	082	-121
88-2	-148	282	208	-110	184	-385	342	-021		-115	333	-039
133	437	-201	059	306	-115	401	-049	051	-115		-305	382
133-1	-139	239	-009	-165	048	-080	073	082	333	-305		037
133-2	216	-061	169	-008	-037	-004	101	-121	-039	382	037	

it casts doubt on the variables involved and is a clear signal to exercise caution in interpreting the variables involved and the factors they define. In general, however, the correspondence between estimated and obtained communalities is satisfactory for a matrix of this size.

The centroid factors shown in Table V were manually rotated to the structure shown in Table VI. The original intent was to reproduce the structure demonstrated by the printed items when analyzed alone, but this was not feasible. Simple structure and meaning were given the priority in determining these rotations.

Table V
Centroid Factor Loadings
(Decimal points omitted)

Item	A	B	C	D	E	F	G	H	h^2
17	314	-239	571	-157	376	-726	-026	-018	1.177
17-1	-244	244	194	-342	522	-413	-017	088	.726
17-2	376	-090	-064	199	131	121	366	006	.359
27	-276	-144	171	-044	210	157	-125	184	.247
27-1	388	-384	-434	-047	-325	-279	-315	-069	.777
27-2	103	109	228	-124	-162	352	169	-428	.451
67	501	-290	-115	164	067	-190	-012	-050	.419
67-1	454	335	-214	-498	323	049	220	-029	.767
67-2	089	443	-300	-274	252	049	195	219	.521
147	438	-165	091	477	239	227	-366	242	.756
147-1	385	-143	302	-105	-350	139	081	232	.473
147-2	-261	214	-239	-257	-178	-119	-035	215	.330
8	443	484	-107	276	-160	-312	-104	010	.651
8-1	-300	-360	-196	-167	040	097	-246	-157	.382
8-2	500	335	-478	479	500	010	-230	049	1.126
48	-074	488	142	217	-085	-158	-213	-253	.453
48-1	180	-095	-433	-090	-046	157	242	084	.329
48-2	110	605	069	092	-118	-094	022	-086	.422
88	328	-510	215	-458	286	084	-045	-282	.794
88-1	-329	327	-185	186	156	323	-161	-079	.445
88-2	125	-439	-349	-179	038	103	-350	-175	.528
133	452	589	203	-179	-197	044	-336	-209	.818
133-1	-134	-098	-504	-240	066	041	099	-182	.388
133-2	270	128	069	-223	-124	197	-209	340	.357

A varimax rotation of these centroid factors was computed and a principal axis solutionwas also rotatated both manually and according to the varimax criterion. The varimax results are difficult to interpret and to relate to the hypotheses, and the principal axis result is no more satisfactory than the centroid one presented here.

Interpretation of Factors: The Printed Items

The structure exhibited by the printed Survey items when they are analyzed separately is clear enough. Seven of the eight are shown to be well suited for their purpose. In the discussion that follows, a single asterisk (*) designates the highest loading exhibited by an item; a double asterisk (**) indicates that the loading is only one greater than .30 shown by that item.

Factor A: Ascendency

Item	Loading	Content
133	.87**	Difficult to get rid of salesmen
8	.66**	Hesitate to ask for exchange or refund
48	.58**	Dislike asking for money or donation
88	.46*	Attempt to get acquainted

All of the four items selected from the Ascendacy scale are here, none of the others are present. Three of the four do not have appreciable loadings on other factors.

Factor B: Restraint

Item	Loading	Content
67	.71**	Makes spur-of-moment decisions
147	.54**	Fond of gambling
17	.51	Prefers lively parties
27	.17	Prefers work requiring attention to detail

This factor is clearly Restraint but has the interesting characteristic that item 27 seems not to be related to the others. It was retained, in the experiment despite its low communality and poor correlation with the factor with the hope that some of the film items would cluster with this one and help define a stronger factor in the larger analysis.

Factor C: Unnamed

Item	Loading	Content
17	-.76*	Preference for lively parties
88	.36	Attempt to get acquainted
27	.26*	Prefers work requiring attention to detail

This factor (at this end of the scale) describes a shy and unassertive person. Its presence is as difficult to account for as its identity. The inclusion of item 17 which presented a troublesome

communality problem in the larger analysis may be a further warning to caution in describing it. This factor also contains the highest loading shown for item 27.

Factor D: Residual

Item	Loading	Content
8	-.328	Hesitates to ask for exchange or refund.

Since it involves only one item, and that with a questionable loading, no attempt is made to name or describe this factor.

Table VI
Rotated Factor Loadings
(Decimal Points Omitted)

Item	A_6	B_2	C_4	D_3	E_4	F_5	G_3	H_3	h^2
17	067	285	358	-252	191	-289	308	-826	1.175
17-1	-609	205	319	-051	289	-123	164	-286	.724
17-2	455	009	073	069	-008	-020	334	171	.359
27	-143	-232	350	086	034	180	-062	-066	.246
27-1	392	-177	-408	022	085	-505	-337	-219	.776
27-2	-225	124	-260	-254	-188	467	002	021	.452
67	461	-059	-099	-169	196	-300	121	-142	.418
67-1	-231	242	-075	-663	329	-204	-044	244	.769
67-2	-359	190	060	-237	223	-219	-023	444	.520
147	658	-077	316	097	440	094	-042	087	.760
147-1	398	157	202	-282	-334	-032	-239	-027	.474
147-2	-376	069	-030	-160	-124	-235	-256	147	.331
8	278	615	-204	150	218	-253	032	145	.653
8-1	-122	-491	-145	223	053	071	-108	-191	.383
8-2	325	132	-097	102	849	-189	189	436	1.125
48	-065	477	201	319	137	230	056	-053	.451
48-1	057	-220	-168	-230	-036	-262	-032	353	.328
48-2	-067	587	-141	085	054	078	047	184	.421
88	089	-298	024	-651	156	086	-085	-485	.795
88-1	-220	-059	-080	307	222	312	-035	381	.446
88-2	249	-471	-263	260	247	-099	-062	-173	.526
133	106	707	-111	-201	245	288	-329	-075	.820
133-1	-286	-337	-354	-119	067	-170	-036	137	.388
133-2	121	176	235	-153	053	-018	-455	152	.358
% CFV	17.3	18.8	8.6	12.34	12.24	9.6	6.4	14.7	

Interpretation of Factors: The Filmed Items

The factors tabulated in Table VI are examined in this section, with references made to the sign of the loading.

Factor A_6 - Rhathymia

Item	Loading	Content
147	.658**	States fondness for gambling
17-1	-.609*	Prefers rowdy parties (male)
67	.461*	Makes spur-of-moment decisions
17-2	.455*	Prefers quiet party (mixed)
147-1	.398*	Gambles for small sums
147-2	-.376*	Gambles for large sums
67-2	-.359	Changes decision in restaurant
27-1	.392	Prefers non-detail work (girl present)
8-2	.325	Throws away defective record

This factor is the opposite of Restraint. The two film items pertaining to parties disagree in content, but the behavior of 27-1 argues against interpretation according to the distraction introduced by the presence of girls; it is equally plausible to specify that most ROTC cadets do not like charades. One of the factors in the varimax rotation resembles this one.

Factor B_2 - Submissiveness

Item	Loading	Content
133	.707*	States difficulty in getting rid of salesmen
8	.615*	Hesitates to ask for exchange or refund
48-2	.587**	Does not want to ask for large contribution
8-1	-.491*	Does not ask for refund for peanuts
48	.477*	Dislikes asking for contributions
88-2	.471*	Turns away without talking to lecturer
133-1	-.337	Buys magazine from street salesman
88	-.298	Hesitates to meet attractive person

This factor is the opposite of Ascendancy; every Survey item is present and each is accompanied by one film item. The film items do not offer support for distracting effect of either extraneous variable, however. It might be inferred that the presence of women in items 88-1 and 133-2 has weakened them as measures of Ascendancy. Like Factor A, this one has a counterpart in the varimax solution.

Factor C_4 - Restraint

Item	Loading	Content
27-1	-.408	Prefers detail work in bookstore (girl present)
133-1	-.354*	Buys magazine from street salesman
27	.350**	Prefers work requiring attention to detail
17	.358	Prefers lively parties
17-1	.319	Prefers quiet party (male)
147	.316	Fond of gambling

The three items from the Restraint scale of the Survey identify this item. All of the film items involve those which used sex as a distractor, so it may tentatively be identified as a secondary sex factor. Item 17-1 disagrees with its parent; this may be evidence that, when they answered the printed item, the respondents were thinking of mixed parties, or that the "lively" party was too vividly represented (see the discussion of assumptions on page 12). Any discussion of item 17 must be made guardedly.

Factor D_3 - Submissiveness

Item	Loading	Content
67-1	-.663*	Changes decision in coffee shop
88	-.651*	Hesitates to meet attractive person
48	.319	Dislikes asking for contributions
88-1	.307	Sits by boy in crowded classroom

Two of these are Survey items which have been shown to be good measures of Ascendancy, and 88-1 agrees with the interpretation they provide. The presence of 67-1, which should have been a Restraint item is not considered to conflict with this interpretation since this item does not appear in any of the factors identified as Restraint.

Factor E_4 - Impulsiveness

Item	Loading	Content
8-2	.849*	Throws away defective record
147	.440	Fond of gambling
147-1	-.334	Does not gamble for small sum
67-1	.329	Makes quick firm decision in coffee shop

This definition must be tentative because: (1) it is based on only four items, (2) three of these are experimental, and (3) one of them, 8-2, has already aroused suspicion because of its communality. On the other hand: (1) the three film items all

refer to quick, but not necessarily well-advised, decisions, and (2) Guilford and Zimmerman (1956) have found evidence that their Restraint factor may not be univocal and that a sub-factor, which they refer to as impulsive in nature, may be present. If this is the case, throwing away the defective record in 8-2 would be consistent.

All of these experimental items were intended to provide the opportunity for distraction by money to appear in the analysis. It is tempting to use this factor as evidence for Hypothesis II, but the other interpretation seems more plausible.

Factor F_5 - Sex Factor I

Item	Loading	Content
27-1	-.505*	Prefers detail work in bookstore (girl absent)
27-2	.467	Prefers non-detail photographic work (girl present)
88-1	.312	Sits by boy in classroom

This is the first plausible evidence that the experimental variables have systematically influenced responses. It might be more impressive if the factor were reflected and the signs and interpretations reversed.

Factor G_3 - Sex Factor II

Item		Content
133-2	-.455**	Walks away from salesgirl in street
27-1	-.337	Perfers detail work in bookstore (girl absent)
133	-.329	Finds no difficulty in getting rid of salesmen
17-2	.334	Prefers quiet party (mixed)
17	.308	Prefers lively parties

This differs from Factor F because it compares film items with the parent Survey items, but never compares two film items with one another. It is regarded as less important than Factor F in supporting Hypothesis II. It involves item 17 and 17-2, both of which are viewed with suspicion (see Table VI and Factor A).

Factor H_3 - Unidentified

Item		Content
17	-.826*	Prefers quiet parties
88	-.485	Hesitates to meet attractive persons
67-2	.444	Makes quick firm decision in restaurant
8-2	.436	Throws away defective record
88-1	.381*	Sits by boy in classroom
48-1	.353**	Asks aggresively for small contributions

The leading item in this factor is one that has been described repeatedly as suspicious; the other items present no clear pattern; therefore, no firm explanation is offered for it. Two conjectures can be offered.

It will be recalled that an unidentified doublet was found in the analysis of the eight printed items; that factor has Survey items 17 (with a large loading) and 88 (with a moderate one). It could be that the factor shown here is the counterpart of that one; if so, it remains unidentified. It may also be noted that, with one exception, these items refer to a retiring and submissive person who seems to be prompted by a desire to avoid personal interactions and to seek the easiest solution to social pressures. Neither of these interpretations is compelling, however, and the factor is not named or explained.

IV. Discussion

Of the eight sets of items presented, each consisting of a verbally stated item from the Survey and its two filmed counterparts, five are found which offer some support for the hypotheses. These will be examined in the same order in which they are found in the preceding text and tables.

Survey item 17 (preference for lively parties) is a troublesome one which complicated the analysis and the interpretation. Its communality in the larger analysis is larger than it should be permitted to be and its interpretation is made difficult for that reason. The two filmed sequences which represent it are poorly concieved; they offer choices between behaviors so extreme that many of the responses may have been made to the less undesirable.

Survey item 27 (preference for work which requires attention to detail) was retained after demonstrating that it was a poor measure of the basic factor in the hope that it would find something in common with some of the film items and reveal its nature. Both of the filmed examples are found in a minor factor (less than ten percent of the common factor variance) where they contradict one another with respect to the basic factor, but agree in the preference indicated for the company of girls. This suggests that a factorially complex item may offer a better possibility of demonstrating the effect of an extraneous element on responses.

Survey item 67 (making spur-of-the-moment decisions) operates well as a measure of Restraint, but the two film items which represent it make no contribution of the basic trait or verification of the hypotheses. The rationale underlying the construction of these sequences seems adequate, but they were not well executed; other extraneous elements, not related to the experimental variables, may have influenced the responses.

Survey item 147 (fond of gambling) appear both as a measure of restraint and of impulsiveness. Neither the verbal nor the filmed items make any contribution to the experiment.

Survey item 8 (returning defective merchandise) and one of its counterparts (the bag of inedible peanuts) are measures of Ascendancy. When a larger sum is involved, as in the film item which depicted a defective item costing five dollars, it becomes a misture of restraint and impulsiveness. It was expected that the phonograph record item would resemble the verbal one, and that the one involving a bag of peanuts would not. The addition of a specific detail, not mentioned in the verbal statement but present in the picture, has altered the factorial nature of the item; this is consistent with the rationale of the Introduction, but operates

in a way that the hypotheses did not foresee. The broken record item (8-2) was one of two in the experiment which demonstrated a communality greater than one, however, so any interpretation of it is suspect.

Survey item 48 (willingness to ask for contributions or donations) functioned in the way that was anticipated. When large donations are mentioned it retains its original effect as a measure of Ascendancy; by making specific reference to small donations, its usefulness for that purpose is destroyed.

Survey item 88 (willingness to make an effort to meet an attractive person) measures both kinds of Aseendancy represented in the analysis. The verbally stated item loads on two factors and is paired with a different filmed item on each. The two kinds of film items measure different traits, but it is not possible to determine the nature of either.

Survey item 133 (getting rid of salesmen) and the film in which a male salesman is represented are measures of Ascendancy. When the unwelcome salesperson is a woman, the only appreciable loading is on one of two sex factors. This conforms to expectation. Like item 8, it demonstrates that factorial content can be altered by making specific in a picture, a detail that was not specific in the verbal statement. Like item 88 it indicates that, when they responded to the original Survey item, these men assumed that the other person involved was a man.

Relationship of Factors to Hypotheses

Hypothesis I stated that the factorial complexity of a set of items presented visually would be greater than that of the same items verbally stated. This is certainlly confirmed; three factors account for nearly all of the variance in the verbal items, while eight factors account for barely half of the variance in the combined analysis.

To confirm the fact of increased complexity is one matter, however; to account for it is another. The structure of the set of 24 items is complex in a way that cannot be completely understood because no reference variables were included in the experiment for the new elements purposely introduced by the film items.

Hypothesis II stated that the nature of the new factors created by the controlled introduction of selected extraneous elements into the pictures could be controlled by the experimenter. Some support is found for this hypothesis, but it is not compellingly clear that the new factors were closely associated with the experimenter's machinations.

Part of that support is in the existence of a factor (B_2) which includes all of the items from the Ascendancy scale of the Survey and just one of the filmed counterparts of each. New variance was introduced by the distracting elements added to the statements, but it appears in combination with one of the basic traits instead of in addition to them. It appears that the factorial nature of a verbal statement may be supported or distorted when extraneous details are added, but predictions are not easily made.

Part of the support is in the existence of two weak factors (F_5 and G_3) both of which have in common the fact that they are defined by items which indicate stable preferences for the company of girls. Survey item 27 performed more satisfactorily than any of the others in this respect.

As an experimental variable, differences in the sex of the persons depicted operated more uniformly and more dependably than differences in amounts of money depicted.

As a vehicle for demonstrating the effects of the variables, the Ascendancy scale is more readily manipulable than the Restraint scale. Experimental contexts which involve visible, rather than implied, differences offer the most promise for future experiments with motion pictures.

V. Conclusions and Recommendations

Like most experiments, this one has yielded information of three kinds: that obtained directly by the ways in which the expectations held for it were realized, that obtained indirectly by the ways in which the outcome departed from those expectations, and that obtained by inference or extension to other contexts.

In the first category, a precedent and a justification for the exploration of motion pictures as testing devices has been provided, evidence has been presented to indicate that questions asked in this way may contain some elements in common with verbal statements but they may also contain some elements specific to the film medium, and a clear warning has been offered to anyone who would indiscriminately substitute one medium for the other.

The second category, that of indirectly obtained information, is a fertile one and is entitled to special consideration in an exploratory experiment because a recognition of the conceptual and operational difficulties which influenced this experiment is the most effective way to improve its successor. Four circumstances involved in planning and executing this experiment can be applied to further research in motion picture testing.

First: An excessive number of limitations were imposed on the specifications for visual content of these film. Future experiments which are not concerned with establishing the fact of departure from a verbally stated item of known content will be freed from such limitations. Judicious selection of basic situations and of experimental variables can provide for sufficient freedom in devising items and scene content.

Second: The importance of access to adequate motion picture production facilities cannot be overemphasized. Adequacy must be evaluated with respect both to technical competency and flexibility.

Technical excellence is required because the subjects for any experiment are sophisticated viewers of commercially produced entertainment films which are essentially perfect technically. Lapses from the accustomed standard will have an effect on the viewer similar to that of grammatical errors and blurred printing on a reader. No reward is attached to the conception of psychometrically clever items if they are beyond the capacity of the production unit to complete.

Third: An assumption was made, forced by the lack of experimental means for evaluating them, that the various instances of the experimental variables depicted would be equally attractive to the subjects. That assumption was clearly not valid. Access to a larger body of experimental precedent in motion picture

testing would make the invalid assumption less damaging; removal of some of the restrictions imposed here on item content would make it easier to fulfill; provision for pretesting and revision would make it unnecessary.

Fourth: Lack of reference materials for the experimental variables hindered the interpretation of this experiment. Visual communication is so specific that the representation of the simplest situation contains an uncounted, but vast, number of elements, some of which will attract the attention of the respondents. Independent measures of the importance of these, if they can be identified in advance by the experimenter, will make precise interpretations easier to find. Single appearances of an element, particularly an unintended one, constitute a source of specific variance in the responses and these can have no effect but to degrade the data.

The third category of information, inferences and extensions of the lessons learned from this experiment, pertains to the direct application of motion pictures to educational testing. The first question proposed here dealt with the possibility that, since they operate in such different ways to accomplish their purpose, words and pictures might not be able to ask the same questions.

From the fragmentary conclusions drawn, it may be concluded that they can ask the same question, but that they can also ask questions which appear to be the same but differ in subtle ways. The test builder who is tempted to turn to motion pictures should ask himself if the qualities he proposes to measure can be found in other ways. If they can, then considerations of cost and mechanical complexity should require that the other method be adopted. If they cannot, then he should assume nothing, equip himself abundantly with funds, patience, and ingenuity, and be willing to accept nothing but the best materials -- conceptually and technically -- that can be produced.

The conclusions and implications drawn from this experiment have been stated in general terms in order to avoid bequeathing its specific concerns and procedures to its successor in the hope that its lessons, but not its shortcomings, may be repeated. Whatever form that successor may assume -- whether it is one suggested here or one not yet concieved -- will demand ingenuity, diligent attention to detail, and a judicious balance between the demands of theory and the limitations of practice. In return, it promises to the fields of communication and measurement, a contribution that cannot be provided by any other medium.

VI. Summary

The difference between words and pictures as communicative devices is discussed; words are characterized by their abstract nature, while pictures refer more directly to a specific attribute or set of attributes of some object or occurrence. The nature of the demands that a test places on a communication system is considered, and the relevance of the distinction between words and pictures to the testing problem is examined.

In the light of the nature of a test and the differences between these two ways of communicating, a question is raised concerning the ability of a verbal statement and a pictorial sequence to ask the same question. An experiment designed to answer that question is described.

Two hypotheses are presented:

I. The factorial complexity of a set of items increases when they are presented to the respondent in the form of pictures instead of words.

II. The nature of the change in factorial content that accompanies their conversion into pictures is controlled by the experimenter through his choice of content of the pictures used.

The vehicle chosen for the experiment was the Guilford-Zimmerman Temperament Survey because it consists of fairly short verbal statements, the scales demonstrate high internal consistency reliability, and the correlations between scales are low.

Four items were chosen from the Restraint scale of the Survey and four from the Ascendancy scale. The factorial structure of this set of eight items was demonstrated. For each item, a pair of short motion picture sequences were prepared, of such a nature that their content might be described by that item.

For two of the items from each scale, one member of the pair of film sequences offers a choice between two kinds of behavior in response to or in the presence of a woman; the other member of each pair depicts a comparable situation in which the central figure is a man.

It is assumed that the subjects differ in the degree of importance they attach to sex differences and in the consistency with which the basic trait is organized in their behavior. The expected effect of the pairing arrangement is an influence on the manner in which the items are percieved by the persons who attach the greatest importance to sex differences and the persons in whom the basic trait is least consistently organized. Instead of resp-

onding to the content of the original question, these persons are expected to respond to visual details present in the picture but not in the verbal statement. This material, which is under the control of the experimenter, is extraneous to the temperament measure but constitutes the independent variable of the experiment.

In the other two items from each scale, the experimental variable was the amount of money depicted or implied in the decision the respondent was required to make. Small amounts were involved in one member of each pair and large amounts in the other. This difference was expected to operate in the same way as sex differences so the distracting effects on temperament measures of two experimental variables could be observed in two contexts.

The anticipated effect was the appearance in the analysis of the original temperament factors defined by the Survey items and some of the film items, and of other factors defined by items which had in common uniformity in the perception of sex or money differences.

The Temperament Survey was administered to the entire freshman class of the Army ROTC at the University of Illinois and, one year later, the film items were administered to the sophomore class. About two hundred cadets provided responses to both forms of the items.

When all twenty-four items were intercorrelated and factor analyzed, the first hypothesis was easily confirmed. Eight factors were found to account for about half the common factor variance in the combined analysis, while four factors accounted for nearly all of the variance in the analysis of the verbal items.

Two of the factors were identified as Restraint, two as Ascendancy, one as Impulsiveness, two were concerned with the effect of sex differences in the responses, and one remained unidentified. Support for the second hypothesis is seen, but it is not strong, nor does it occur in the way in which it was expected.

There is good evidence that one member of each pair of film items made from the Ascendancy scale is, itself, a measure of that trait and that the other member of that pair is not. This is evidence that the experimental variables have performed their function of influencing responses to the pictorially stated questions. The differences between pairs in which sex is the extraneous influence is stronger and more uniform than differences when amount of money is the experimental variable.

It was concluded that motion pictures offer promise as a testing medium, but that their use may be more complex than was originally supposed, and that a larger body of experimental precedent be accumulated before definitive assertions are made.

References

Carpenter,C.R., Greenhill,L.P., Hittinger,W.F., McCoy,E.P., McIntyre,C.J., Murnin,J.A., and Watkins,R.W., Development of a sound motion picture proficiency test, Personnel Psychology, 1954, 7, 509-523.

Gibson,J.J., Motion Picture Testing & Research, AAF Aviation Psychology Program, Report No. 7, Washington, D.C., Government Printing Office, 1947.

Guilford,J.P. and Zimmerman,W.S., The Guilford-Zimmerman Temperament Survey, Manual of Instructions and Interpretation, Beverly Hills, Calif., Sheridan Supply Co., 1949.

Guilford,J.P. and Zimmerman,W.S., Fourteen dimensions of temperament, Psychological Monographs, No. 417, 1956, 70, Washington, D.C., The American Psychological Association.

Johnson,D.M. and Vogtmann,W.G., A motion picture test of achievement in psychology, American Psychologist, 1955, 10, 69-71.

Lhotsky, Jaromie, Film test and test films, in Speer,E. (Ed.) Die Vortrage der 4 Lindauer Psychotherapiwoche, 1953, 113-117

Lhotsky, Jaromie, Film test and test films, in Psychological Abstracts, 1955, 29, No. 7293.

McIntyre, C.J. Sex, age, and iconicity as factors in projective film tests, Journal of Consulting Psychology, 1954, 18, 337-343.

Rose, Nicholas, et al, A report of the development of certain motion picture tests of motivation and personality, Unpublished report of research performed under Contract AF 18 (600)-349 between the University of Southern California and the Human Resources Research Center, 1953.

Spiegelman, Marvin, Evaluation of personality by viewing a motion picture, Journal of Projective Techniques, 1956, 20, 212-215.

Stoller, Robert and Geertsma,R.H., Construction of a final examination to assess clinical judgment in psychiatry, Journal of Medical Education, 1958, 33, 837-840.

FROM:

ERIC FACILITY
SUITE 601
1735 EYE STREET, N. W.
WASHINGTON, D.C. 20006