Behavior Research Methods & Instrumentation
1974, Vol. 6, No. 2, 111-121

SESSION III

SYMPOSIUM: COMPUTER DATA GENERATING

DANA B. MAIN, *University of Michigan,* **Presider**

Generalizing the problem definition step in the computer simulation of factorial experiments

CHARLES E. HALLENBECK and ROBERT B. WELCH
University of Kansas, Lawrence, Kansas 66045

A simulation system (KUSIM) used at the University of Kansas is described and examples provided. The system permits instructors and students without computer skills to bring computer simulation of psychological experiments into the classroom.

There is a popular cosmology which holds that at some time in the distant past our universe came into existence as the result of a single creative act by a supreme being. Since that time, according to this view, it has been incumbent upon Man to explore that universe by formulating suitable questions to "put to Nature." Well-formulated questions are rewarded by yielding reliable knowledge. Badly formulated ones lead to confusion, uncertainty, and all the other evils.

Two views of education emerge from this cosmic adventure. One is that man becomes educated as he learns more about the universe in which he lives. The other view is that man becomes educated as he learns how to ask better questions or design better experiments. The first view is content directed; the second is method directed.

Inspired by Richard Johnson's (1973) "DATACALL Game," last year at the University of Kansas we undertook the task of implementing certain aspects of the cosmic events described above. We reasoned that if an instructor were given tools for "creating" a psychological universe, and if his students were given tools for conducting "probes" into that universe in the form of experiments, then perhaps something of educational value for both participants could be realized. Our system therefore evolved into two main-line FORTRAN programs which share a common library of subprograms, and more important, which share a common data file. The first of our programs is called "GODSIM," and it has access to the data file in both reading and writing modes. The second program is "MANSIM," and it has read-only access to the file.

THREE COMPONENTS OF DATA-GENERATING SYSTEMS

As an aid in comparing our system to other data-generating systems, we have identified three major aspects of such systems that are of importance. These are: (1) appearance to the problem authors, (2) appearance to the problem users, and (3) internal data structures. We believe that our system exhibits important differences from other schemes on at least the first and third points, perhaps to a lesser degree on the second point.

Appearance to the Problem Authors

Our GODSIM program implements a context-free, or Chomsky Type 2, grammar (Berztiss, 1971, p. 164ff.) for the processing of problem descriptions supplied by instructors who wish to "play God" and author a new problem universe. The language is simple and easy to use. The grammar is defined, and examples of statements are given in "Formal Descriptions of Researchable Problems" (see Appendix A). Essentially, the author supplies a problem title and subtitle on a "DEFINITION" statement. He then supplies from 1 to 12 "FACTOR" statements which add structure to the problem universe. Each factor has a fixed number of levels, and mnemonic identifiers for the factors and their levels are supplied. At this point, the scope of the problem universe is specified, but it is barren of effects. Any experiments performed on such a problem universe will yield only unit normal deviates with mean zero and sigma one. The problem author completes his problem definition by supplying optional "MEAN," "ERROR," and "EFFECT" statements.

The "EFFECT" statement permits the association of values to the various levels of factors. It also permits values to be associated with interacting factors to any depth of complexity. If desirable, a problem with 12 factors could incorporate an interaction of Order 12. Short-cut methods are available for describing lengthy strings of numbers.

Appearance to Problem Users

The "MANSIM" program implements a similar context-free, or Chomsky Type 2, grammar and processes a second simple language for use by students in the description of experiments. The grammar is defined and examples of statements are given in "Formal Descriptions of Factorial Experiments" (see Appendix B). The student supplies a "USE" statement containing the problem title he wishes to use, a "FOR" statement containing his own name, and an "EXPERIMENT" statement containing an identifying number. That number resets internal conditions in the program and helps the student identify his output when he receives it. He then proceeds to describe the experiment, group by group. He does so by supplying "SET" statements, each containing the key word "SET" followed by the name (or symbol) of the factor to be set and the name (or number) of the level to which it should be set. Factors not set are assigned levels at random. Such random settings are not reported or discoverable by the student. When a group description is complete, a "RUN" statement containing the sample size is given. The "MANSIM" program reports simulated data in maximally compact format, to a precision which is 10% of the problem's residual error. If greater or lesser precision is desired, the student experimenter can override this output format by supplying a "FORMAT" statement of his own in each group definition. Ordinary FORTRAN rules prevail, except that "I" and "F" conversions are equally available options to the student E.

Internal Data Structures

The common data file shared by "GODSIM" and "MANSIM" serves as a repository for the information associated with all problems included in the system. Our computer system permits device-independent files, and so the number of problems we can operate on simultaneously is essentially unlimited. Search times through such a file will undoubtedly mount as the number of implemented simulations increases, but such considerations have not yet become significant in our use of the system. The important point concerning our internal data structures is that they do not constitute mathematical modeling in the usual sense of that term. Authors of problem universes do not explicitly develop a mathematical model of the problem area which they are creating. The absence of a formal mathematical model is perhaps a strength as well as a weakness. It is a weakness because it fails to provide problem authors a strong conceptual tool. It is a strength because it allows an empirical description of the effects of a set of independent variables on a single dependent variable. These effects are stored and retrieved for direct use by student Es.

The data structures in the common file which correspond to a given problem are simple FORTRAN arrays. Some of these contain label information for the factors and their levels; their structure is trivial. However, one array contains the information from which effects are reconstructed. This array is the complete Cartesian product of factor levels. Its cells contain the mean value of the dependent variable under the indicated combination of independent variables. This structure, too, is simple. The problem author need not supply cell values for all table positions, since they can easily be generated if not specified. The most detailed table which the problem author must supply is that for the highest order interaction effect he wishes to specify.

In selecting our data structures, we have adopted a literal interpretation of a remark by Richard Johnson (1973), who conceptualizes the DATACALL procedure as an "analysis of variance in reverse." We have therefore implemented an analysis of variance table with cell values representing the population parameters rather than the sample statistics. We have exercised a "tradeoff" between required storage space and generality. Our approach is costly on the first criterion but achieves considerable generality.

CAPABILITIES AND LIMITATIONS

There are significant limitations to our present implementation of the "GODSIM-MANSIM" programs. Some of these are enumerated here. First, all problems are univariate. Only a single dependent variable exists in a given problem. We are currently enlarging this area to include up to four dependent variables. To define such problems, one set of "FACTOR" statements will be used, while a complete set of "MEAN," "ERROR," and "EFFECT" statements will be used for each dependent variable. Furthermore, within-cell correlations between all six pairs of the four variables, as well as identifying label information, will require new statement types for the "GODSIM" language. We anticipate extending the "MANSIM" language by adding an "OBSERVE" statement to specify which of the dependent variables should be observed in a given experiment or run.

Second, the current implementation does not permit "repeated measures" or "within-Ss" designs. While the implementation of multiple dependent variables would take an initial step toward this capability, a more direct approach would be desirable. We have in mind extending the "GODSIM" language to permit test-retest reliability and "rules" governing the repeatability of measures within Ss to be added tothe data structures. The form of such extensions is not yet clear. The "MANSIM" program would have its language extended by including a "VARY" statement within the group definitions, indicating the factor or factors to be varied within Ss. Practice effects would also need to be considered in the problem definition.

Third, all independent variables are categorical in nature and have default attributes of randomness. Each factor is either set at a specific level or is drawn randomly from among its available levels, with equal probability of each level being drawn. The result of this limitation is that "SET" statements must be given for two types of independent variables: those controlled in a given experiment and those being manipulated. We would like to enlarge the language by including a "DISTRIBUTE" statement as an option for each "FACTOR" statement, so that unequal probability of random selection can be associated with the various levels of a factor. In one current problem, the factor "HAIR-STYLE" has three levels: "LONG-HAIR," "SHORT-HAIR," and "BALD." Random selection means that when "HAIR-STYLE" is not controlled or manipulated, approximately one-third of every sample drawn is bald.

Fourth, the dependent variable has a normal distribution with uniform variance over all factors and with no upper or lower bounds to its possible values. We are hard pressed to design realistic problems in which the dependent variable is percentage information. We occasionally obtain negative percentages and percentages in excess of 100%. We are remedying this situation by extending the "GODSIM" language to include "FLOOR" and "CEILING" statements. These statements, in addition to having a minimum and maximum value, will be modified by the terms "ACCEPT," "REPORT," or "IGNORE," indicating the action to be taken when "bad data" are generated. For "ACCEPT," the floor value or ceiling value will be supplied in place of the illegal value. For "REPORT," the program will discard the illegal value and get a replacement which is within range. It will add a summary of bad data counts for the E. For "IGNORE," the action will be the same as for "REPORT" except that there will be no indication to the E that illegal data were encountered while generating acceptable values.

Ideally, we would like to permit several common transformations to be applied to the output data after their generation, but have not pursued this extension very far. Such transformations would permit nonhomogeneous variances to arise in response to certain experimental designs.

CLASSROOM EXPERIENCE WITH "KUSIM"

In addition to requirements for a major in psychology, our department recommends a certain sequence of courses for students wishing to apply for graduate training. That sequence consists of a 3-h course entitled "Statistics in Psychological Research," taken in the sophomore year, followed by one or more of our four laboratory courses in experimental psychology, each of which is a 6-h course. The four courses are entitled "Experimental Psychology: Human Learning," "Experimental Psychology: Perceptual and Response Processes," "Experimental Psychology: Social Behavior," and "Experimental Psychology: Personality." To date, we have had experience with "KUSIM" in two of these courses, "Experimental Psychology: Perceptual and Response Processes" and "Experimental Psychology: Social Behavior," on a basis which has not permitted its systematic evaluation.

Neither of the instructors who have used "KUSIM" has used it as a substitute for actual experimental work, but as a supplement. Both agree that no program, no matter how many options or facilities it contains, is adequate to achieve the entire educational task. Student response, while favorable, has depended critically on a careful preparation of what Johnson calls a "scenario." That is, a clear and complete description of the research area, identifying its importance, exploring some unanswered questions, and outlining at least some of the available factors to be explored. An example of such a scenario is provided in "Research Scenario: IQ Scores" (see Appendix C).

One of our two courses has employed the simulation embedded in a game situation, with costs and payoffs (see "A Computer Game," Appendix D), while the other has not. Although in neither case are we able to present the results of systematic pre- and postexposure measures of skills or attitudes, it is possible to give a general description of our experience with the game-playing situation. Nearly all of the students in the course were favorable toward their experience and felt that the experiment simulation game should be retained for use in future classes. There were, of course, some problems. Possibly the most serious of these was the failure to set aside occasional class periods devoted to a discussion of the "experiments" that had been run up to that point and the corresponding results. This procedure, which is easily implemented, would serve both to reduce unintended "replications" and to provide the students with a partial picture of the "research universe," thereby increasing their chances of doing profitable future experiments. A second problem was the use of a cost-payoff structure which provided insufficient "profit" for the discovery of a main effect. In order to make the game most involving, it is probably important that the "rewards" be worth the effort. A further way to make the game more interesting (and more like the "real world" of research) is to avoid the use of fixed costs and payoffs. Rather, students could be "charged" different amounts for the manipulation (or control) of different factors and could receive payoffs which vary in size as a function of how theoretically important or unexpected the statistically significant effect is judged to be. Such judgments could be made either by the class or by the instructor.

At the University of Kansas, we are at least 2 years behind the current state of affairs at the University of Michigan (Main, 1971, 1972; Main & Head, 1971), with respect to the degree of commitment to the use of computer simulation and to the evaluation of its effects. We do not have large experimental courses under the directorship of a single faculty supervisor. Therefore, we must resort to a colleague-to-colleague "sales job" with respect to computer simulation. Our initial experience has been fairly successful, but we are not on the verge of a computer revolution.

SUMMARY

The University of Kansas is using a simulation system, "KUSIM," which permits instructors and students without computer skills to bring computer simulation of psychological experiments into their classrooms. It provides convenient free-form languages for the definition of arbitrary researchable problems and for the specification of the factorial experiments to be performed. Two independent main-line programs and a large library of subprograms are written in Honeywell's "FORTRAN Y," a super-FORTRAN. The program that performs the creative act of establishing the psychological problem is named "GODSIM." The other program has "read-only access" to the GODSIM file and simulates experiments on problems obtained from it. It is called "MANSIM."

Problems may have from 1 to 12 factors. Each factor is named and given a number of levels. Each level is also named. The manifold product of levels over all factors may not exceed 4,096, and their sum may not exceed 512. Any individual factor, or any interacting set of them, may be given an "effect." Factors may be defined without any effect. Grand means and residual error are also specified.

The user names the problem to be studied and then describes each group of Ss to be "run." Each group is described by naming the factors to be set, the levels at which to set them, and by giving the number of Ss. The program produces "raw data" and identifies the problem, the user, the factor and level names, and the sample sizes.

The system has implemented only the simplest model of factorial experiments. It does not permit repeated measures designs or error distribution other than normal. While its most imaginative application is in a game-playing environment, it does not implement any payoff structure or cost evaluation of simulated experiments. It does, however, illustrate the ease with which problems can be defined and retained for later use.

APPENDIX A

FORMAL DESCRIPTIONS OF RESEARCHABLE PROBLEMS

CHARLES E. HALLENBECK

A simple context-free grammar is described, and its application to the description of researchable problems in psychology is illustrated. The language of this grammar is processed by a FORTRAN "Y" program. A data structure corresponding to the problem description is added to a file for later retrieval. A second FORTRAN "Y" program processes formal descriptions of factorial experiments that might be performed on the problems. Simulated raw data are produced by applying the experiment descriptions to the problem descriptions.

In general, a context-free grammar consists of a set of expressions called "productions," which consist of a "nonterminal" on the left side and a mixture of terminal and nonterminal symbols on the right. The productions are "permissive" rules for rewriting the nonterminals. A valid statement in the language described by such a grammar is one which contains only terminal symbols, but which can be derived from, or reduced to, the "distinguished nonterminal symbol" of the grammar.

Notation

In addition to upper- and lowercase alphabetic symbols, the productions use numeric constants, angle brackets, square brackets, parentheses, and the right arrow. The following rules govern the use of these symbols:

(1) Uppercase alphabetic symbols are literals and are punched or typed just as they appear.

(2) Symbols enclosed in angle brackets are replaced by information supplied by the user as he writes statements in the language.

(3) Symbols enclosed in square brackets are optional and may be omitted if desired.

(4) Symbols enclosed in parentheses, with no qualifying prefix before the left parenthesis, must appear at least once and may be repeated an indefinite number of times.

(5) Bracketed symbols enclosed in parentheses, with no qualifying prefix before the left parenthesis, are optional and, when present, may be repeated an indefinite number of times.

(6) Symbols enclosed in parentheses, with a numeric prefix before the left parenthesis, must be repeated as many times as indicated by the prefix number.

(7) Bracketed symbols enclosed in parentheses, with a numeric prefix before the left parenthesis, are optional and if present may be repeated as many times as indicated by the prefix number.

(8) The right arrow is used to separate a nonterminal symbol on the left side of a production from the expression forming the right side. The arrow is understood to mean "may be rewritten as."

Var1 → (Var2) FINISH

Var2 → DEFINITION <Problem Title> <Problem Subtitle> [Var3]

Var3 → Var4 11 ([Var4]) [Var5]

Var4 → FACTOR <Letter> <Factor Name> <Number> Nr(<Level Name>)

Var5 → [Var6] [Var7] ([Var8])

Var6 → MEAN <Value>

Var7 → ERROR <Value>

Var8 → EFFECT <Letters> Nr(<Value>)

Fig. 1. A context-free grammar for "GODSIM."

Example 1
Removing a preexisting problem from the GODSIM file.

DEFINITION TEST-TAKING ANXIETY

PROFESSOR A. L. JONES

FINISH

The problem title begins with the first nonblank character following the key word "DEFINITION." It contains 24 characters, including imbedded blanks. The subtitle is a similar 24-character field, starting with the first nonblank character on a new line following the major title.

Example 2
Defining a new problem having one factor with three levels with that factor having no effect on the dependent variable.

DEFINITION TEST-TAKING ANXIETY

PROFESSOR A. L. JONES

FACTOR A AGE 3 YOUNG MIDDLE-AGED ELDERLY

FINISH

The presence of a "FACTOR" statement causes the named problem to be defined and added to the file. The letter symbol "A" following the key word "FACTOR" is required. Factor and level names may be up to 24 characters, but may not contain imbedded blanks. The number following the factor name must agree with the number of level names that follow. New lines may be started at any point should more space be needed.

Example 3
Defining a problem with two factors, one with three levels and the other with two, with only the second factor having an effect on the dependent variable.

DEFINITION TEST-TAKING ANXIETY

PROFESSOR A. L. JONES

FACTOR A AGE 3 YOUNG MIDDLE-AGED ELDERLY

FACTOR B SEX 2 MALE FEMALE

EFFECT B 20, 30

FINISH

A second factor is introduced, and it must then use the letter "B" after the key word "FACTOR." To give this factor an effect, reference must be made in an "EFFECT" statement to the symbol "B." Since that factor has two levels, two values must be described for the factor, one for each level.

Example 4
Defining a two-factor problem in which only the interaction of the two factors has an effect on the dependent variable.

DEFINITION TEST-TAKING ANXIETY

PROFESSOR A. L. JONES

FACTOR A AGE 3 YOUNG MIDDLE-AGED ELDERLY

FACTOR B SEX 2 MALE FEMALE

EFFECT AB 20, 30, 25, 25, 30, 20

FINISH

The "EFFECT" statement associates an effect with the "AB interaction" by supplying six values for the various cells of that interaction in lexicographic order (i.e., A1 B1, A1 B2, A2 B1, A2 B2, A3 B1, and A3 B2).

Example 5
Defining a problem with two factors, in which a grand mean and error term are specified, as well as an interaction effect.

DEFINITION TEST-TAKING ANXIETY

PROFESSOR A. L. JONES

FACTOR A AGE 3 YOUNG MIDDLE-AGED ELDERLY

FACTOR B SEX 2 MALE FEMALE

MEAN 25

ERROR 5.0

EFFECT AB -5, 5, , , 5, -5

FINISH

The "MEAN" statement establishes a grand mean of 25 units for the dependent variable, and a residual error of 5.0 units is also specified. Absence of these statements always implies a zero mean and residual error of unity. The "EFFECT" statement now employs two "null" values, and these are calculated automatically so as to make the six cell values total zero. The cell values are increments over and above the grand mean.

APPENDIX B

FORMAL DESCRIPTIONS OF FACTORIAL EXPERIMENTS

CHARLES E. HALLENBECK

The formal description of a researchable problem is converted by a FORTRAN "Y" program into a data structure for storage and later retrieval. It is then possible to supply formal descriptions of factorial experiments to be performed on those problems. This paper presents a context-free grammar for formal descriptions of factorial experiments and illustrates their use to achieve simulation of research in psychology. A second FORTRAN "Y" program processes the descriptions of experiments and generates simulated raw data by applying the experiment description to the problem description. Random number generators guarantee the same "Type I" and "Type II" errors in the output of the simulated experiments as exist in the real world.

Notation

In a companion paper to this one, a notation is given for context-free grammars. That notation is employed here as well. The terms used in that notation are supplemented by a slash, used between elements which are alternatives. One or the other of two elements separated by a slash, but not both, may be selected.

Var1 → (Var2) STOP

Var2 → USE <Problem Title> (Var3)

Var3 → FOR <Name of User> (Var4)

Var4 → EXPERIMENT <Identifying Number> (Var5)

Var5 → [Var6] ([Var7]) RUN <Sample Size>

Var6 → FORMAT <Layout Specifications as in "FORTRAN">

Var7 → SET Var8 Var9

Var8 → <Letter> / <Factor Name>

Var9 → <Number> / <Level Name>

Fig. 1. A context-free grammar for "MANSIM."

Example 1

Performing a one-group experiment with no controlled variables.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 1
RUN 100
STOP
```

One hundred Ss will be "run" in a simulated experiment. However, since no factors were specifically mentioned, all factors existing in the problem definition will be drawn at random from their possible levels.

Example 2

Performing a one-group simulated experiment with one controlled and one uncontrolled factor.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 25
SET AGE YOUNG
RUN 25
STOP
```

Twenty-five Ss are run with "AGE" controlled at level "YOUNG"; and the other variable, not mentioned, randomly drawn from its various levels.

Example 3

Simulating a two-group experiment with one controlled and one uncontrolled variable.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 6
SET SEX MALE
RUN 30
SET SEX FEMALE
RUN 30
STOP
```

Two groups of Ss are run, with 30 males in Group 1 and 30 females in Group 2.

Example 4

Simulating a 2 by 3 factorial experiment, with six groups of Ss.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 12
SET AGE YOUNG
SET SEX MALE
RUN 10
SET AGE YOUNG
SET SEX FEMALE
RUN 10
SET AGE MIDDLE-AGED
SET B 1
RUN 10
SET A 2
SET B 2
RUN 10
SET A ELDERLY
SET SEX 1
RUN 10
SET AGE 3
SET B FEMALE
RUN 10
STOP
```

Six groups are run, with 10 Ss in each group. The data produced by this simulated experiment would be suitable for a 3 by 2 factorial analysis of variance, but the choice of statistics is left to Henry Gibson. Notice that factors and levels may be referred to in "SET" statements by their proper names or by letter symbol and level number, as preferred. In any event, output contains complete labeling by name and symbol, no matter how specification was achieved.

Example 5

Simulating two experiments on the same problem for the same user, a three-group experiment and a two-group experiment, each controlling one factor and randomly assigning the other.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 101
SET AGE YOUNG
RUN 20
SET AGE MIDDLE-AGED
RUN 20
SET AGE ELDERLY
RUN 20
EXPERIMENT 102
SET SEX 1
RUN 30
SET B FEMALE
RUN 30
STOP
```

The first of these experiments tests 60 Ss in three age groups, within 20 Ss per group. Each group consists of a random mixture of males and females, since sex was not controlled. The second experiment also tests 60 Ss in two groups of 30, one group of males and one of females, with their ages randomly distributed over the three possible levels.

Example 6
Two experiments performed by different users on the same problem, one with a request for special output format.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 50
FORMAT (5X, 3F12.4)
SET AGE YOUNG
RUN 45
FORMAT (10X, 5I12)
SET AGE ELDERLY
RUN 45
FOR SUSY SMITH
EXPERIMENT 60
SET SEX MALE
RUN 4
SET SEX FEMALE
RUN 16
STOP
```

Henry Gibson gets his data in "F" or "I" formats, which he imposes on the output. Susy Smith, however, permits the program to format her output. She has asked for data from 4 males but 16 females, for reasons known only to her.

Example 7
Two experiments on different problems performed by the same user.

```
USE TEST-TAKING ANXIETY
FOR HENRY GIBSON
EXPERIMENT 3
SET AGE MIDDLE-AGED
RUN 30
SET AGE ELDERLY
RUN 20
USE LONG TERM MEMORY
FOR HENRY GIBSON
SET UNCERTAINTY HIGH
SET LANGUAGE ENGLISH
RUN 300
SET UNCERTAINTY HIGH
SET LANGUAGE FRENCH
RUN 300
STOP
```

The appearance of a second "USE" statement necessitates that Henry identify himself again. The file will be searched for the two independent problems, "Test-Taking Anxiety" and, later, "Long-Term Memory." The nature of the dependent variables in these experiments is assumed to be known to the user and is not described in the output.

APPENDIX C

RESEARCH SCENARIO: IQ SCORES

CHARLES E. HALLENBECK

Perhaps one of the most intensively studied human characteristics is "general intelligence." At first the attention of psychologists was drawn to the problem of making the classification of school children as objective as possible. Later, the requirements of military manpower classification and industrial selection made the concept important with adults as well.

While the general intelligence of adults has been widely studied, there is still the need to refine the notion of IQ and make its measurement as reliable, as valid, and as efficient as possible. Imagine that you have offered a new IQ test to meet the need for better measurement, and that you want to examine the way in which your new measure relates to various characteristics of adults. You can select as many adult Ss as you wish for the purpose, although you have only five characteristics to consider in making your selection.

First, age is a consideration. If you do not specify age as a criterion for selection, your Ss will be drawn at random from three age groups of adults: those less than 20 years of age, those from 20 to 50 years of age, and those over 50 years of age. You may select Ss to fall into any of these three groups if you believe that age should be studied in relation to its effect on IQ scores.

Second, the sex of your Ss is a criterion which you may or may not want to study. You may draw your Ss from a population of all males or all females, or you may choose to ignore sex. In that case, your Ss will be drawn at random from available males and females.

Third, the nature of the occupation which your Ss engage in constitutes a factor in which you might be interested. Only a rough index of this variable is available to you, i.e., whether their occupation is "white-collar" or "blue-collar." If you do not choose to study this variable, your Ss will be drawn randomly from the two categories.

Fourth, your Ss can be selected according to a personality characteristic which has become popular among many researchers today. That characteristic is called "locus of control," and your Ss are classified as having either "external" or "internal" locus of control. That means that they believe their success or happiness depends on fate or factors outside their own control (such Ss are "externals"), or that they are the masters of their lives and can determine their own success or happiness (such Ss are "internals"). Once again, by ignoring this factor, your Ss will be a random mixture of the two types of locus of control.

The fifth and final factor in which you might be interested is a rather unusual one. It may have significance to some persons interested in the outcome of your research, or it may not. In any event, it is easy to measure, and so it is included for your consideration. That factor is the hair style of your Ss. There are three types of hair style for you to consider: long hair, short hair, and bald. While baldness may not seem to be a hair style at first glance, it is grouped together with the other two categories to provide three different types of hair style for purposes of your research. While baldness may be quite rare in the general population, it will be no more or less difficult for you to obtain bald Ss as either of the other categories. Consequently, if you ignore this factor, your Ss will be drawn randomly from these three equally likely categories.

Your task, therefore, is to try to discover how your new measure of adult IQ depends on any or all of these five factors: age, sex, occupation, locus of control, or hair style. A good IQ test should be relatively insensitive to some of those factors, but might properly reflect some true differences in general intelligence associated with other factors. See what you can learn.

APPENDIX D

A COMPUTER GAME

ROBERT B. WELCH

The Computer Game is designed to give you, the student, an opportunity to carry out simulated experiments on a specified problem by means of a computer program. The problem that was chosen is the effect that repeated exposure to the Mueller-Lyer (M-L) illusion has upon the strength of the illusion. Typically, the effect is a *decline* in strength, hence the problem is entitled "M-L illusion decrement." On the basis of actual research literature, your instructor has chosen a number of variables, some or all of which affect the amount of decrement experienced after a fixed number of exposure trials. The actual experimental setup and nature of the independent variables and dependent variable (decrement of illusion) will be described in more detail below. The computer has been programmed such that certain levels of certain variables have specific effects, again based on the results of actual experiments reported in the literature. Note again that not all the variables, nor all the levels of a given variable, necessarily have effects upon the dependent variable.

Basically, the game is to run simulated experiments, using one or more of the specified independent variables in an attempt to discover the nature of the "universe" created by your instructor. You carry out an experiment merely by specifying (on a special form) the independent variable(s) and levels that you wish to use and the sample size (N) per group. Your teaching assistant will use this form in order to "run the experiment" on the computer. The "results" will be received quite soon (possibly within the hour) in the form of a score for every S (according to group), the sum of the scores, and the sum of the squared scores. In order to see if your variable(s) has (have) had any effects upon the dependent variable, you will have to statistically analyze the data (probably by means of analysis of variance). Finally, you must *very briefly* (1-2 pages) report the design of your experiment (the variables manipulated, sample size, etc.), your summarized data, the results of the statistical analysis, and your conclusions. After doing this, you can begin a new experiment.

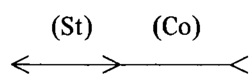
One of the reasons why this exercise is called a "game" is that you are playing for points. There are certain costs for certain designs (e.g., it costs more points to run a two-factor experiment than a one-factor experiment) and certain payoffs for certain outcomes (e.g., you get a *big* payoff for each statistically significant effect that you find and a *small* payoff for failing to find an effect). Your goal, of course, is to amass as many points as possible.

While this exercise is admittedly (and purposefully) a game, it has a number of aspects to it that resemble the "real world" of research activity. Clearly, the presence of costs and payoffs makes the game analogous to the real situation. That is, it costs money and time to run experiments, and some experimental designs are more expensive than others. Furthermore, one receives payoffs that vary from small to large as a function of the outcome of the experiment. These payoffs are rarely, if ever, directly monetary, but include publications, prestige, promotions, and, of course, the satisfaction of intellectual curiosity. The main part of the game that is unlike actual research activities is that, with the computer, an experiment can be run in a matter of minutes. If you actually tried to set up and run the same experiment on real Ss, it would take anywhere from 2 to 6 months. Thus, one great advantage of the Computer Game is that it allows you to run many experiments in a very short period of time. This makes it possible for you to use the results of one experiment as the basis for the design of a new experiment. If you are thoughtful in your designs and perceptive in examining your results, it should be possible for you to describe the "universe" of the present research area by running only 6-8 experiments (although you certainly are not limited to this number).

AN INTRODUCTION TO THE RESEARCH PROBLEM

(1) The Experimental Paradigm

As previously indicated, the topic chosen as the basis of the Computer Game is the decremental effect that repeated exposure to the Mueller-Lyer (M-L) illusion has upon the strength of the illusion. The illusion looks like this:



The two horizontal lines are actually equal in length, but because of the "tails," Co looks longer. A typical experimental situation is one in which the S is first measured on his initial experience of the illusion. The "Method of Adjustment" is often used, in which S adjusts the length of the horizontal line in the comparison (Co) portion of the figure until it appears to be equal in length to the horizontal line in the standard (St) portion. Because when St and Co are *objectively* equal in length, Co *appears* to be longer, and therefore S will *reduce* the length of Co in order to make it *appear* equal to St. The precise amount by which S shortens Co is measured to the nearest millimeter. The larger the error, the stronger the illusion S must have experienced. Several trials are taken in order to establish a reasonably stable measure of the *initial* strength of the illusion. At no time is S provided with any feedback regarding his performance. After this "before" measure has been taken, S is given repeated exposure to the illusion. Each exposure is of fixed length, and there is an interval of time between each. S is not required to adjust Co and is not given any information to suggest that he is viewing an illusory stimulus object. After a fixed number of exposures (e.g., 100), an "after" measure of the strength of the illusion is taken in the same manner as the "before" measure. *The dependent variable is the difference between the strength of the illusion before the exposure period and its strength afterward.* If the difference indicates a *reduction* in the strength of the illusion, the number is given a "+"; if the illusion has actually *increased* in strength (a very unlikely event), the number is given a "−." A control group should (ideally) also be run, in which S is given "before" and "after" measures of the strength of the illusion, but during the exposure period presented with some nonillusory figure. You should assume that this group, if run, would reveal *no change* in the strength of the illusions. Thus, the results of this control group will be ignored hereafter.

(2) The independent variables

(A) Factor A: Exposure/Trial
Levels 1-7: .2, .4, .8, 1.6, 3.2, 6.4, 12.8 (sec)

This independent variable is the amount of time (in seconds) that the illusion is exposed on each of the 100 trials. A tachistoscope is used in order to guarantee that the exposure time per trial, for a given condition, is precise and reliable.

(B) Factor B: Prior-Knowledge
Levels 1-2: Yes, No

This variable represents whether or not S has heard of and/or seen the M-L illusion *before* being in the experiment. Strictly speaking, this does not qualify as an independent variable because it is not actually *manipulated* by the investigator.

(C) Factor C: Figure-Orientation
Levels 1-2: Same, Rotated

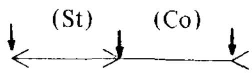
This independent variable is whether the M-L illusion figure is presented in the same orientation (e.g., St on the left and Co on the right) on all 100 exposure trials or, instead, is rotated on every trial (on half the trials St is on the left and Co is on the right, and on the remainder it is the reverse).

(D) Factor D: Figure-Size
Levels 1-3: Large, Medium, Small

It is possible for the figure to be of three different sizes long, medium, or small. A "large" figure is 12 in. long; a "medium" figure is 6 in. long; a "small" figure is 3 in. long. In all three cases, the distance from S's eyes to the figure (in the tachistoscope) is 1.5 ft.

(E) Factor E: Eye-Movement
Levels 1-3: Fixate, Free-Saccadic, Directed-Saccadic

This independent variable involves what S does with his eyes during each exposure to the M-L illusion figure. In the "fixate" condition, S fixates the point in the middle of the figure where St and Co meet. In the "free-saccadic" condition, S is allowed to move his eyes over the figure in any manner he wishes. In the "directed-saccadic" condition, S is required to deliberately move his eyes between three different points (denoted by the arrows in the following figure):



(F) Factor F: Practice-Distribution
Levels 1-3: Massed, Distributed, Very Distributed

This variable represents the duration of the interval of time between each of the exposure trials. In the "massed" condition, the interval is 2 sec, in the "distributed" condition, it is 10 sec. In the "very distributed" condition, it is 20 sec.

(G) Factor G: Sex
Levels 1-2: Male, Female

The operational definition of this (obviously nonmanipulable) variable is self-evident.

(H) Factor I: Field-Dependence
Levels 1-2: Field-Dependent, Field-Independent

This S variable is measured by means of the Rod-and-Frame Test (Witkin). In this task, S faces a luminous rectangular frame in the center of which is a luminous rod. Except for these objects, the room is completely darkened. The frame is placed on a *tilt* (28 deg right or 28 deg left) and the rod is also on a tilt (either in the same or opposite direction as the frame). The task is to set the rod (via a switch) to *true* (gravitational) upright. A person is designated "field-dependent" if the tilted frame strongly influences his setting of the rod (i.e., S sets the rod on a tilt in the same direction as the frame-tilt). A "field-independent" S is one who is quite accurate in setting the rod to true vertical (i.e., he is not influenced by the frame).

HOW TO PLAY THE GAME

(1) Your first step is to examine the nature of the research problem and the variables which you can use. Consider which variables (and which levels of a given variable) you think are likely to affect decrement of the illusion.

(2) "Design" an experiment to test the variable(s) you think is (are) likely to have an effect. Decide which variables (if any) you wish to hold constant (note that the effects of any variable which you neither manipulate nor control will be distributed randomly among your Ss). Decide upon the same size (N) you wish to use for each group. Remember–each of these decisions is associated with a certain cost (which will be spelled out below).

(3) Once you have made your decisions, fill out *two* copies of the "Decision Sheet" (see below) and give them to your instructor or teaching assistant. He will actually "run" your experiment by means of the teletypewriter at the computer terminal.

(4) The "results" of your experiment will be returned to you as quickly as possible (probably in a day or two from submission). The second copy will be kept by your instructor.

(5) Next, you must statistically analyze your data, in order to see if you have found any effects. The experiment is set up as a *random groups* design. Thus, the range of appropriate statistical tests is relatively limited. The simplest one that you could use would be a *t test*–if you were to manipulate only one factor and used only two levels of it. If you were to use three or more levels of one factor, the appropriate statistical test would be a *one-way analysis of variance*, together with a post hoc test of differences between groups means. If you were to use two independent variables, the appropriate statistic would be a two-way analysis of variance.

(6) Finally, you must write a *very short* report (e.g., 1-2 pages) of your experiment. It should include (a) a brief description of the design (i.e., which factor(s) manipulated at which levels, using what N), (b) a table with the group means, (c) results of the statistical analysis (including the necessary calculations), and (d) your conclusions.

(7) You should have calculated the cost of your experiment before actually running it. The payoff will be calculated by your teaching assistant and he will keep you apprised of your current "financial status." No payoff is recorded until you have "published" your paper.

(8) After noting the results of your first experiment (and how many points you now have), you should design a new experiment. Presumably, you will now have a little better idea of the nature of the research universe and can therefore be more confident that your new design will garner a significant number of points. It is also possible to *replicate* your experiment (or somone else's, for that matter), but the payoff for this is significantly reduced (see "Rules of the Game").

(9) You may run as many experiments as you wish (or can afford). However, you are only *required* to publish *five* of them. Be sure to keep each of your Decision Sheets so that you have a record of all your experiments and their payoffs.

(10) Note: It is very important that you be aware of the results of the experiments run by the other members of the class. In this way, you will avoid replications (unless you are replicating on purpose). Your instructor and teaching assistant will make "public" (in a manner to be decided upon) the results of all "published" experiments. You may also wish to talk to other members of the class about the results of any "*unpublished*" studies which they may have run.

RULES OF THE GAME

You will be "given" 200 points at the outset of the game. Your task, of course, is to use your points wisely and to attempt to increase your total. It is permissible to "go into debt" (on the hopes of getting a large enough payoff to get back "into the black").

(1) Costs

(A) Number of Ss per group:

N per group	5	10	15	20	25	30	35	40	45	50	60	75	90	100
Costs per group	14	17	24	30	35	40	44	48	51	54	59	65	70	72

Any number of Ss can be chosen per group, up to 4,500 individuals. The cost is a negatively accelerated function of the number chosen. The costs for Ns up to 100 are given above.

(B) Number of factors tested:

No. of factors tested	1	2	3	4	5	6	7	8
Cost	10	20	40	70	110	160	210	290

In the "real world" of research, some factors are more expensive to examine than others. However, for the purposes of the game, we will ascribe an equal cost to each of the eight factors; the more factors you examine in a given experiment, the more expensive it is.

(C) Number of factors controlled:

No. of factors controlled	1	2	3	4	5	6	7
Cost	5	10	14	17	19	20	20

In the "real world," some factors are more expensive to control than others. However, here we shall make them all equally costly. Note that the maximum number of variables that can be controlled is seven, since at least one of the variables is being tested.

(2) Payoffs

(A) Main effects:
  (1) Statistically significant
    (A) "Original"–80 pts. each
    (B) Replication–20 pts. each
    (C) Second replication–0 pts.
  (2) Statistically nonsignificant
    (A) "Original"–20 pts. each
    (B) Replication–0 pts.

(B) Interactions:
  (1) Statistically significant
    (A) "Original"–30 pts. each
    (B) Replication–10 pts. each
    (C) Second replication–0 pts.
  (2) Statistically nonsignificant
    (A) "Original"–10 pts. each
    (B) Replication–0 pts.

(C) Confounded experiments–0 pts.

REFERENCES

Berztiss, A. T. *Data structures: Theory and practice.* New York: Academic Press, 1971.

Johnson, R. DATACALL: What it is and how to do it. Unpublished paper, Exxon Education Foundation, New York, New York, 1973.

Main, D. A computer simulation approach for teaching experimental design. Paper presented to the American Psychological Association, Washington, D.C., 1971.

Main, D. Toward a future-oriented curriculum. American Psychologist, 1972, 27, 245-248.

Main, D., & Head, S. Computer simulations in the elementary psychology laboratory. Paper presented at the Conference on Computers in the Undergraduate Curricula, Dartmouth College, June 1971.

Behavior Research Methods & Instrumentation
1974, Vol. 6, No. 2, 121-123

Modeling and the Michigan Experimental Simulation Supervisor: An overview and some prospects

ROBERT L. STOUT
Mathematical Psychology Program, University of Michigan, Ann Arbor, Michigan 48104

A data-generating system (MESS) is described which facilitates the construction and study of behavioral science models of various types. The system provides a spectrum of options that make it possible to employ MESS in a variety of classroom situations with relatively unsophisticated students.

The Michigan Experimental Simulation Supervisor (MESS) is a program written in IBM FORTRAN IV (G); its basic purpose is to facilitate the construction and study of behavioral science models of all kinds. One version or another of MESS (formerly known as Expersim) is now in use at more than 20 universities, colleges, and community colleges in the U.S. and other countries in undergraduate and graduate courses in experimental design, statistics, and several content areas of psychology.

When MESS was designed, five principal considerations were used in making decisions concerning the program specifications: (1) It should be possible for students with all degrees of lack of sophistication to learn to use the system quickly; and students should spend as much time as possible on the design and analysis of their experiments and as little time as possible puzzling over how to get the computer to do them. In particular, the program should (a) require as few lines of input as possible, (b) allow students to describe conditions and groups using terms derived directly from the language of the problem area, (c) be highly tolerant of minor errors in spelling and syntax, (d) provide helpful error messages, and (e) provide output which is formatted and labeled in such a way as to be maximally intelligible. (2) The system should be capable of handling models from any area of psychology, and of any desired structure or degree of complexity. The program should also allow a large number of independent and dependent variables, including nonnumeric dependent variables, and every kind of model structure, from static analysis of variance and regression models to dynamic, highly structured models such as finite automata or cognitive models like Newell and Simon's General Problem Solver. (3) The system should allow any general class of experimental design, including multivariate designs, repeated measures, confounded designs, and correlational experiments. (4) The system should provide a command language with as many options as possible to facilitate use of the system and provide pedagogical flexibility. (5) It should be as simple as possible to implement and modify models in the system.

The versions of MESS that have been produced so far have all been strongly student-oriented; that is, extraordinary effort was expended to make sure that