DOCUMENT RESUME

ED 389 993 CG 026 738

| | |
|---|---|
| AUTHOR | O'Connell, Ann Aileen; Bol, Linda |
| TITLE | Development of a Tutoring System for Probability Problem-Solving. |
| PUB DATE | [95] |
| NOTE | 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995). |
| PUB TYPE | Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160) |
| EDRS PRICE | MF01/PC01 Plus Postage. |
| DESCRIPTORS | College Students; *Computer Assisted Instruction; Higher Education; *Intelligent Tutoring Systems; Mathematical Concepts; *Probability; *Problem Solving; Programmed Tutoring; Statistical Analysis; *Statistics; Tutorial Programs; Tutoring |

ABSTRACT

Interpreting information regarding health risks, crime statistics, and government polls requires some ability to use and interpret probabilities. Studies have shown that even after training or coursework in probability and statistics, people still have many difficulties solving probability problems. The thesis of this document is that helping students develop more efficient schema for solving probability problems will improve their performance on such problems, subsequently leading to a stronger and more lasting ability to use and interpret probabilistic information normatively. A proposed seven-step instructional model provides a description of steps involved and knowledge required for successful solution of many different types of probability problems. A computerized tutoring system which incorporates the model, investigates its success as an instructional aid in teaching and learning elementary probability concepts and procedures. A formative evaluation of the system was undertaken, but is not reported in this document. The development of the system, a description of the evaluation process, and future goals are reported. Additional information includes a table of type and frequency of observed errors in probability problem solving (n=50); a copy of the form used in evaluation, and a visual example of the computer program. (JBJ)

# Development of a Tutoring System for Probability Problem-Solving

by Ann Aileen O'Connell and Linda Bol

University of Memphis

# Development of a Tutoring System for Probability Problem-Solving

Ann Aileen O'Connell
Linda Bol
University of Memphis

Paper presented at the Annual Meeting of the American Educational Research Association
April 18-22: San Francisco, CA.

## Introduction

An understanding of basic concepts regarding probability is essential in our society. Interpreting information regarding health risks, crime statistics, government polls, the likelihood of natural disasters such as earthquakes or hurricanes, all require some ability to use and interpret probabilities. Extensive research documents the existence of biases in people's reasoning about probability and probabilistic events (Tversky and Kahneman, 1974, 1983; Garfield and Ahlgren, 1988; Konold, 1989). Studies have also shown that even after training or completion of a formal course in probability and statistics, people still have many difficulties solving problems that require the use of probability or in making informed decisions under uncertainty (see, for example: Shaughnessy, 1981; Tversky and Kahneman, 1983). Therefore, preparing students for their future in education, business, the social sciences, biology, etc., and enabling them to become informed citizens and consumers would be facilitated if the process of teaching and learning basic probability concepts and procedures was better understood.

The current project is a continuation of previous work in the area of probability problem-solving, which offered an instructional model for how people typically work towards successful solution during probability problem-solving (O'Connell, 1993a; 1993b). That research also indicated that nearly 23% of observable errors in probability problem-solving performance are due to errors in text comprehension, with an additional 45% due to procedural errors. However, many procedural errors result from a misunderstanding of text information. Our present project is based on the belief that helping students develop more efficient schema for solving probability problems would improve their performance on such problems, subsequently leading to a stronger and more lasting ability to use and interpret probabilistic information normatively. We argue that guiding people in the skills required to organize problem information accurately is important to the development of a reliable schema for probability problem-solving.

Reasoning during problem-solving is highly dependent on the specific type of problem a subject is working through, since some problems are by nature more difficult than others and may require qualitatively different kinds of knowledge. The proposed instructional model provides a description of the steps involved and the knowledge required for successful solution of many different types of probability problems. This seven-step model identifies key areas where probability problem-solving is likely to go wrong. The seven steps used in this model and incorporated into our tutoring system are described below:

*1. Understand the given information.*
In order to understand the information provided and develop a representation of the problem, or to interpret the given information as a mathematical or probabilistic expression, the student needs to have adequate knowledge of the natural language of probability, as well as an understanding of the concept of probability itself.

*2. Identify what is being asked (the goal).*
Identifying the goal statement involves the ability to translate the question being asked into a probability statement suitable for solution to the problem. The student needs to distinguish between the following possibilities: is the question asking for a numerical solution to a problem, for the verification of an assumption, or for the verification of a particular formula (i.e., for conditional probability)? In addition, the student must know the meaning of 'at least', 'at most', 'no more than', etc.

*3. Develop notation for the given information and the goal statement.*
This step requires successful completion of steps one and two above, as well as an understanding of the formal, symbolic language of probability in terms of events being described as sets of outcomes. The

student must correctly develop a notation for expressing the given information and the relationship between the given information and the goal.

*4. Identify the correct sample space for the problem.*
The student needs to review possible assumptions or the state of events given in the problem, such as: are the events equally likely or not, are the events independent, are they mutually exclusive? This requires some real world knowledge about events (cards, coins, elevators that work independently, electrical components in series, people's opinions given independently, etc.). In addition, the student requires an understanding of when to assume that a concept (equally likely, mutually exclusive, independent) holds, and how to verify that it is true, if necessary.

*5. Select a method of solution.*
There may be many different methods of solution appropriate for any particular problem. Successful solution rests on choosing an appropriate method, such as using equations only, using tree diagrams, contingency tables, or Venn diagrams. Recognizing problem types will facilitate the choice of a particular method of solution. Therefore, recall of a method of solution used for a similar problem has an important impact on problem solving ability.

*6. Computing the solution.*
The procedure chosen for determining the solution depends on both the problem involved and the solution method decided upon. Generally, solution methods fall into four categories: using equations only; the use of tree diagrams; contingency tables; and/or Venn diagrams. Successful solution also depends on the ability to switch to a different method if the first does not offer a helpful path towards the desired goal. Occasionally, problems may be solved with a combination of these four methods. Each of these solution styles requires slightly different knowledge for normative use.

*7. Is the solution reasonable?*
Evaluating the feasibility of a solution is one of the most important steps in successful probability problem-solving. It requires real world knowledge, and also an appreciation for the basic tenets of probability theory, i.e., that probability is never negative or larger than one. The solution found for any particular problem should 'make sense' to the problem-solver.

## PURPOSE OF THE PROJECT

The purpose of this research project is to investigate the success of the above model as an instructional aid in teaching and learning elementary probability concepts and procedures. To this end, the model has been incorporated into a computerized tutoring system. A formative evaluation of the system is currently being conducted to ascertain how well the tutor actually provides the student with the knowledge and skills required for solving the types of problems typically encountered in a first course in probability and statistics. The formative evaluation will provide a test of the model as well as of the tutoring system, and modifications to the system or the model will be based on the results of our evaluation.

## DEVELOPMENT OF THE TUTORING SYSTEM

This project is currently funded through a Faculty Research Grant to the first author (University of Memphis). As recommended by other instructional design teams, we utilize a multidisciplinary approach to the planning and design process, emphasizing different areas of expertise (Morrison & Ross, 1988). Our development team includes two graduate assistants in instructional design, a content expert (O'Connell), and an evaluation expert (Bol).

The HyperCard system was chosen as the platform for the probability tutor, primarily due to the availability of Macintosh computers at most schools and colleges. HyperCard offers many advantages as an instructional tool, most specifically in its linking capabilities. Our software offers students the ability to jump from unit to unit, based on their needs as well as to their responses to practice and test questions supplied throughout the program.

The system contains five instructional units. Each unit includes a test and review section and a summary card detailing the major points covered. The first three units focus on introductory information regarding probability, and the last two units incorporate our instructional model. The units are identified as follows:

1. Understanding Probability
2. Working with Sample Spaces
3. Probabilities of Outcomes and Events
4. Probabilities for Compound Events
5. Dependence and Conditional Probability

Through the linking capabilities of HyperCard, a student can choose which of the topics to be studied, and can review the summary cards for any of the units at any point in time. Question cards posed throughout the tutor include a "hint" and an "answer" button to provide the student with optional assistance. The student inputs their answer(s) to a question or a series of questions directly onto the card. If an error is made, the tutor provides the student with information regarding the nature of the error. Error identification is based on previous work by the first author (O'Connell, 1993a; 1993b). Table 1 describes the types of errors typically found among students studying probability for the first time. A more detailed description of each error is available from the first author.

Four types of errors are listed in Table 1. These include text comprehension errors, conceptual and procedural errors, and errors in arithmetic. By incorporating both the instructional model and the error analysis information, the tutoring system provides a unique opportunity for tailored instruction. Several cards from Unit 4 are provided in the Appendix.

## EVALUATION PROCESS

Currently, we are in stage one of our planned four stage evaluation. The subject matter expert (O'Connell) reviews a print copy of each HyperCard unit for content accuracy, usefulness of the graphics, typographical errors, and overall flow of each unit and the system. Recommended changes to improve each unit are completed by the instructional designers and the process is repeated until the final version for each unit is obtained.

Stage two of our evaluation will involve obtaining student feedback on each of the five units. Student volunteers will be recruited from introductory statistics courses at the University of Memphis and their reaction to the probability tutor will be assessed. Form 1 will be used to ascertain the clarity of material presented, ease of use, and how strongly the reviewers felt the tutor was successful in conveying the information for each unit. Based on this information, recommended changes will be made to the tutoring system. The capability of the tutor to monitor student performance will also be evaluated at this time, and recommended changes will be incorporated into the tutor.

The plan for stage three of our evaluation involves combining all probability units and evaluating the entire system. At this point we will also evaluate the successfulness of our 7-step instructional model. This will be done by comparing results of a paper-and-pencil test for those students who utilized the tutoring system versus students who were not exposed to the system. The assessment will be conducted in two ways: (a) in terms of per cent correct, between the students utilizing the probability tutor and those students who were not exposed to the tutoring system and instructional model, and (b) the types of errors made by students in both groups will be categorized according to the classification scheme developed in previous research (O'Connell, 1993a; 1993b), and the incidence of particular errors will be compared across both groups of students.

Stage four of our evaluation process will involve upgrading the tutoring system as necessary, and field testing among students in different settings (schools, colleges, universities, and/or majors).

## THE FUTURE

Our overall hypothesis is that those students who were exposed to the instructional model through the use of the probability tutor will be more successful at solving problems involving probabilities and understanding of probability in general, compared to students who were did not use the tutor. Additionally, it is anticipated that students tutored in the instructional model will exhibit fewer types of text comprehension and procedural errors than those in the non-tutored group.

This work provides a significant contribution to the study of teaching and learning, particularly in the area of diagnostic assessment. Understanding why particular patterns of errors occur and where they occur most frequently during the problem-solving process can help us provide more meaningful instruction, as well as strengthen learning theories in this and other domains.

Table 1
Type and frequency of observed errors in probability problem-solving (n=50 students)

Text Comprehension Errors

| Type | Label | Freq. | % of T |
|---|---|---|---|
| T1 | Missassigning stated probability value | 53 | 38.4 |
| T2 | Incorrect specification of goal (equality) | 13 | 9.4 |
| T3 | Choosing pairs instead of triples/singles, etc. | 0 | 0 |
| T4 | Misinterpretations of inequalities | 16 | 11.6 |
| T5 | Selection with vs. without replacement | 2 | 1.4 |
| T6 | Real world knowledge errors | 1 | 0.7 |
| T7 | Incorrect model of experiment described in problem | 42 | 30.4 |
| T8 | Interference from another (previous) problem | 11 | 7.8 |
| Total | | 138 | (23.1)* |

Conceptual Errors

| Type | Label | Freq. | % of C |
|---|---|---|---|
| C1 | Misconceptions: defn. of probability/sample space/n(S) | 0 | 0 |
| C2 | Misconceptions: frequency vs. probability | 2 | 1.8 |
| C3 | $p>1.0$ | 11 | 10 |
| C4 | $p<0$ | 0 | 0 |
| C5 | $P(S)\neq 1.0$ | 0 | 0 |
| C6 | formal language of probability | 7 | 6.4 |
| C7 | Misconceptions: equally likely events | 63 | 57.3 |
| C8 | Misconceptions: mutually exclusive events | 17 | 15.5 |
| C9 | Misconceptions: independence | 4 | 3.6 |
| C10 | Misconceptions: mutually exclusive vs. independence | 5 | 4.5 |
| C11 | Misconceptions: complementary events | 1 | 0.1 |
| Total | | 110 | (18.4)* |

Procedural Errors

| Type | Label | Freq. | % of P |
|---|---|---|---|
| P1 | Procedural errors in determining sample/event space | 9 | 3.3 |
| P2 | Incomplete/unfinished | 19 | 7.0 |
| P3 | General use of formulas | 11 | 4.1 |
| P4 | Procedural errors involving independence | 96 | 35.4 |
| P5 | Procedural errors involving mutual exclusiveness | 27 | 10.0 |
| P6 | Procedural errors involving sequential experiments | 6 | 2.2 |
| P7 | Procedural errors involving use of tabled data | 45 | 16.6 |
| P8 | Procedural errors involving conditional probability | 34 | 12.5 |
| P9 | Procedural errors involving complementary events | 11 | 4.1 |
| P10 | Inventing incorrect procedures or rules | 13 | 4.8 |
| Total | | 271 | (45.4)* |

Arithmetic Errors

| Type | Label | Freq. | % |
|---|---|---|---|
| A | Totals: Arithmetic errors | 54 | (9.0)* |

Unclassified Errors

| Type | Label | Freq. | % |
|---|---|---|---|
| X | Totals: Unclassified errors | 34 | (5.7)* |

* Percent of total errors (total=597)

Form I: Volunteer Participant Evaluation Form

**Unit Reviewing:** ________________________________

## FORMATIVE EVALUATION

Please respond to each of the statements listed below by circling the appropriate response you feel matches your feeling regarding the probability tutor instructional unit.

**SA = STRONGLY AGREE**
**A = AGREE**
**NF = NO FEELING**
**D = DISAGREE**
**SD = STRONGLY DISAGREE**

| | | | | | | |
|---|---|---|---|---|---|---|
| 1. | I felt that the unit was informative. | SA | A | NF | D | SD |
| 2. | The amount of material covered was about right. | SA | A | NF | D | SD |
| 3. | I felt the unit was easy to understand. | SA | A | NF | D | SD |
| 4. | I felt comfortable with my ability to move through the unit. | SA | A | NF | D | SD |
| 5. | I now have a good picture in mind of the concepts presented in this unit. | SA | A | NF | D | SD |
| 6. | Overall, the unit was successful in conveying information | SA | A | NF | D | SD |
| 7. | After completing this unit, I felt more knowledgeable about the topic the unit covered. | SA | A | NF | D | SD |

Please provide any additional comments about the unit or the tutoring system itself:

________________________________________

________________________________________

________________________________________

________________________________________

________________________________________

Your Name: ____________________ Date: ____________

Please Indicate by Checking:

Major Field of Study/Work:
Education Research ________
Instructional Design ________
Math/Statistics ________
Other ________ (write in, please) ____________

Garfield, J., and Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. Journal for Research in Mathematics Education, 19, 1, 44-63.

Konold, C. (1989). Informal conceptions of probability. Cognition and Instruction, 6(1), 59-98.

Morrison, G. R., and Ross, S. M. (1988). A four-stage model for planning computer-based instruction. Journal of Instructional Development, 11, 6-14.

O'Connell, A. A. (1993a). A classification of student errors in probability problem-solving. Unpublished doctoral dissertation, Teachers College, Columbia University, New York.

O'Connell, A. A. (1993b). "Investigating the relationship between conceptual and procedural errors in the domain of probability problem-solving." Presented at the Annual Meeting of the Mid-South Education Research Association, November: New Orleans, LA.

Shaughnessy, J. M. (1981). Misconceptions of probability: From systematic errors to systematic experiments and decisions. In Shulte, A. P. and Smart, R. J. (Eds.), Teaching Statistics and Probability (pp. 90-100), NCTM 1981 Yearbook. Reston, Virginia: National Council of Teachers of Mathematics.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90(4), 293-315.

## Compound Events

The probability that a child in a certain school district has either the chicken pox or the flu is .58. A nurse practitioner has determined that, for this district, the probability that a child has only the flu is .18, while for chicken pox this probability is .60. What is the probability that a child has both?

First, let's try and define the information that we are given.

We, then, will be able to solve this problem step by step.

Menu

Exit

## Compound Events

1. What events are we given information about?

   chicken pox
   flu
   either chicken pox or flu

2. What are we being asked? What is the goal of the problem?

   Find the probability of a child having both chicken pox and the flu.

Menu Question Card

Exit

## Compound Events

3. How can we represent this information using probabilities?

Given Information: P(chicken pox) = .60
P(flu) = .18
P(either chicken pox or flu) = .58

Goal: P(chicken pox and flu) = ?

Click on the question button to go back to the question card to make sure you can find this information in the problem.

Menu | Question Card | Exit

## Compound Events

What rule or rules should we use to solve this problem? Do we have all the information we need to use a chosen rule?

Review these rules and choose the one you think might help to solve this problem.

**Rule 1**: P(A and B) = 0 if mutually exclusive

**Rule 2**: Addition Rule for M.E. event (when P(A and B)=C).
P(A or B) = P(A) + P(B)

Yes | No

More Rules on next card ...

Menu | Ques. | Exit

## Compound Events

**Rule 2**: General addition rule

P(A or B) = P(A) + P(B) - P(A <u>and</u> B)

or its re-expression

P(A and B) = P(A) + P(B) - P(A or B)

Yes No

**Rule 3**: For Complementary Events

Yes No

P(not A) = 1- P(A)

Menu Ques. Back Exit

## Compound Events

5. Good Choice! We can use our "re-expression" of the addition rule, since we have all the other information we need to solve the problem.

P(A and B) = P(A) + P(B) - P(A or B)

P(pox and flu) = P(pox) + P(flu) - P(pox or flu)
= .60 + .18 - .58 = .20

So we find that the probability that a child has both the chicken pox and the flu is .20.

Does this answer "make sense" to you?

Menu Ques. Exit