

THE EFFECTS OF EVALUATION MODELS,
DATA TECHNIQUES, AND OCCUPATION
ON PERCEPTIONS AND UTILIZATION OF
EVALUATION TECHNIQUES

By

BARBARA ELLEN HOWELL
ff

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1983

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 1986

Thesis
1986
H859e
cop. 2



THE EFFECTS OF EVALUATION MODELS,
DATA TECHNIQUES, AND OCCUPATION
ON PERCEPTIONS AND UTILIZATION OF
EVALUATION TECHNIQUES

Thesis Approved:

Dianna L. Newman

Thesis Adviser

Joseph Pearl

Robert W. Perry

Norman N. Durbin

Dean of the Graduate College

1251274

ACKNOWLEDGEMENTS

The researcher wishes to express appreciation to all who have contributed toward the completion of this study, particularly to the writer's committee: Dr. Dianna L. Newman, Chairman; Dr. Joseph Pearl; and Dr. Katye Perry, for their time and effort.

Special gratitude is expressed to Dr. William D. Frazier, for his encouragement and advise in completion of the study, and for the privilege to work under his direction in the Research Coordinating Unit at the State Department of Vocational and Technical Education.

Sincere thanks is also extended to the writer's family for assistance and encouragement; to my Mother, my brothers, Charles and Gary, my sister, Mary Ann, and their families. It is to her family that the writer dedicates this study.

TABLE OF CONTENTS

Chapter	Page
I. ARTICLE	1
Theoretical Background	
Methodology	8
Participants	8
Vignette	9
Instrumentation	11
Results	11
Discussion and Conclusions	14
References	19
II. EXTENDED REVIEW OF LITERATURE	25
Introduction	25
Statement of the Problem	27
Review of Literature	27
Models in Evaluation	28
Data Techniques	30
Audience Needs in Program Evaluation	33
Statement of Hypotheses	37
III. EXTENDED METHODOLOGY	38
Introduction	38
Sample	38
Vignette	40
Instrument	42
Experimental Design	45
Procedure	48
BIBLIOGRAPHY	50
APPENDIXES	56
APPENDIX A - ADVOCACY WORDING QUESTIONNAIRE	57
APPENDIX B - ADVERSARY WORDING QUESTIONNAIRE	72

LIST OF TABLES

Table	Page
I. Means and Standard Deviations	23
II. Test of Model, Level of Data, Occupation and their Interactions	24
III. Alpha Reliability Coefficients for the Four Dependent Constructs	46

LIST OF FIGURES

Figure	Page
1. Experimental Design	47

CHAPTER I

ARTICLE

Evaluation is an important aspect of education. Many decisions are made concerning the implementation of programs by use of evaluation techniques (Rutman, 1982). This utilization of evaluation may vary according to the approach used for evaluation and its purpose. Evaluation approaches differ depending upon the communication model and level of data implemented when describing a program to a particular audience.

One way of examining communication is through the communication paradigm (Fishbein & Ajzen, 1975). This model suggests that "who says what to whom" has an effect on how information is received. The message source, the medium by which it is transmitted, and to whom it is being transmitted, all play an important role in influencing the decisions to be made (McGuire, 1969). Hawkins, Roffman, and Osborne (1978), have found different media and modes to be useful in creating an effective communication model, thereby identifying procedures which will gain ultimate utilization of evaluation results.

The main purpose of searching for an effective communication model in evaluation is to aid in the presentation of information and feedback. One way to do this is to provide the audience with a suggested utilization of the evaluation information (Alkin, 1980). This may be accomplished by use of differing evaluation methods. Information provided to a particular audience may also differ with the amount of data presented to reinforce the evaluator's presentations. For example, Brown and Newman (1982) found that increased levels of sophistication of data effect use of information. The purpose of this study is to examine the relationship between the type of data presented, the type of evaluation method used, and the professional level of the audience and their effect on utilization of evaluation. A hierarchy of data received will be varied among administrators and teachers within one of two evaluation wordings: advocacy and adversary. The construct of concern in studying these variables will be their effect on: 1) respondents' perceived agreement with the evaluators recommendations; 2) respondents' perception of the information provided; 3) respondents' utilization of information; and 4) respondents' perception of the overall presentation of evaluation information.

THEORETICAL BACKGROUND

Research has shown that different models of evaluation do indeed initiate different outcomes, and are and should be prepared and presented differently (Worthen & Sanders, 1973). A plethora of evaluation designs have been developed to meet variations which may effect the decision-making process. In this particular study, two evaluation designs were developed that varied by wording of recommendation: advocacy wording and adversary wording. These wordings reflect components of a larger model, the judicial evaluation method (Wolf, 1979). In this method, decision-makers are presented with recommendations, typically based on the same information, that reveals both the positive and negative aspects of the program. At the advocacy end, the evaluator presents recommendations in a positive or 'pro' manner. At the opposite end of the judicial spectrum lie the adversary recommendations which are presented to the audience in a negative or 'con' manner (Kourilsky, 1973). In studying the judicial model, Brown and Newman (1982), found that decision-making in an evaluation context was influenced by the order in which the arguments (pro-con as compared to con-pro) were presented in an evaluation. When the adversary

argument was presented first followed by the advocacy argument (con-pro), audience support was more in favor of recommendations than when the advocacy argument was presented followed by the adversary argument (pro-con).

A question not yet answered is what is the effect of wording, advocacy or advocacy, when presented alone? If the presence, and hence the order, of both models affects perceptions of evaluation, it may be that positive recommendations alone are more acceptable than are recommendation, phased in a negative manner, but supported by the same evidence. Brown, Braskamp, and Newman (1978) found working of reports to significantly affect perceptions of reports. Reports loaded with professional jargon were rated as more technical and more difficult than were reports written in the lay language. Newman and Howell (1985) found that educators rated evaluation reports which used negative recommendations as more objective than those using positive reports.

A second variable related to use of evaluation is the degree of sophistication of data presented in support of evaluative recommendations. In a study by Brown, Braskamp, and Newman (1978), it was found that the use of data interacted with the language used in educators' perceptions of an evaluation report.

In this study language was defined as being either jargon-loaded or jargon-free. Jargon-loaded reports were those that used professionally derived terms to succinctly convey a concept to a particular group, whereas jargon-free reports used lay language. These styles of language were crossed with two variations in empirical presentation. Objective reports were defined as those using percentages to justify findings. Subjective reports presented no percentages or any other form of data to substantiate evaluator opinions. The results of this study showed that jargon-loaded subjective reports were rated by the audience as more technical than were jargon-loaded objective reports, and were found to be the most difficult type of report to read. The report rated as easiest for the audience to read and understand was the jargon-free subjective report.

In a second study where the operational definition of data was expanded beyond subjective and objective (Brown & Newman, 1982), four different variations in data styles were examined and found to influence decision-making. These variations included: 1) no data; 2) percentages; 3) percentages and graphs; and 4) percentages, graphs, and inferential statistics.

Results showed that educators agreed more with the evaluator's recommendations when the data presented were in the form of percentages and graphs. When an inferential statement was included, the level of agreement was lower than any other style of data usage, even that of the no data report. In a later study, hypotheses concerning response to recommendations were also tested for the same four data presentations (Brooks, 1984). In this study, however, no significant differences were found in response to the recommendations of the program, ratings of the evaluator, and evaluation information as a result of variations in data presentations.

A third area of variation in the evaluation process is the type of audience involved in the decision-making process. House (1978) indicates that no matter which audience is involved, perceptions of evaluation and its usefulness are key elements in improving utilization. Two prominent educational audience's are administrators and teachers. A study by Thompson and King (1981) found administrators to be more attentive toward evaluators' technical ability than were teachers when considering evaluation information. Administrators were also found to be more attentive to the technical merits contained within a report than were teachers.

Professional level, professional area, and gender are other audience characteristics found to be related to client reactions. A study by Brooks (1984) found a significant difference in both the response to the recommendations and ratings of the evaluator and report as a result of varying organizational roles (teachers, principals, administrators, and evaluators). Evaluators rated recommendations made in an evaluation significantly lower than did the other levels of organizational roles. The interaction of audiences' professional background and the evaluators' title were found to be related to ratings of evaluator credibility (Braskamp, Brown, & Newman, 1978). Administrators tended to rate evaluations as more useful in a decision-making role than did teachers. Also, when the evaluator was introduced as a "researcher" opposed to an "evaluator" or "educator", reports were considered to be more acceptable. Newman, Brown and Littman (1979) indicated that an audience consisting of professionals were more critical of an evaluator than was a novice audience. This supports Carter's (1971) proposition that the closer the audience is to the decision-makers role, the more critical it is of the evaluator. This study is concerned with differences in perception of utilization

of evaluation between teachers and administrators when receiving varying presentations of evaluations.

Methodology

Participants

A random sample of 167 teachers and 156 administrators within vocational education across the state of Oklahoma served as the two audiences for this study. Sixty-seven percent of the sample analyzed were male while 33 percent were female. The range of the participants' age was 24 to 64 years. The median age of the subjects in the advocacy wording model was 40, and the adversary wording model had a median age of 42. Participants were randomly selected from data files available with the State Department of Vocational and Technical Education. Teachers were defined as those currently teaching a vocational education course. Administrators were chosen from those principals who were located at a high school containing one or more vocational education programs. The first sample of randomly selected teachers and administrators was chosen in April, 1985 (n =400; return rate=57 percent; final return n=226). To increase power, another random sample was drawn the following October (n=200; return rate=48 percent; final return n=96). An examination of the nonrespondents

indicated no major differences from respondents for gender, geographic location, or area of teaching. In addition, no differences on the same variables were found for the April and October samples, thus the two groups were combined. Each participant was randomly assigned to a vignette containing one of the evaluation models (Pro, Con), with a particular level of data (No Data, Percentage Data, Percentage and Graphical Data, or Percentage, Graphical and Inferential Statistics) included in the simulation.

Vignette

The vignette in this study is a description of a pilot program utilizing a computer assisted guidance program which had been tried in several schools. The simulated evaluation included information about the teachers' reactions, projected costs of the program, the amount of training required for implementation, and how similar student data were being used before the program was implemented. These simulations and recommendations were based on those validated in previous research by a panel of administrators and teachers who indicated that the settings were typical of decisions in education (Brown & Newman, 1982).

Half of the vignettes presented the recommendations using the advocacy (pro) wording while the other half used the adversary (con) wording. In the advocacy approach, the evaluator presented arguments which were in support of the major recommendations, i.e. "All schools with more than 250 students should use the computerized pupil guidance program". In the adversary approach the evaluator presented arguments which were against the major recommendations, i.e. "Schools must have at least 250 or more students before using the computerized pupil guidance program".

Crossed with each of these models of evaluation were four different data conditions: (1) No Data, where recommendations were not reinforced with any data which might help in the decision-making; (2) Percentage Data, where percentages were used to reinforce the recommendations; (3) Percentage and Graphical Data, where not only percentages, but also a graph depicting those percentages were used to supplement the recommendations; and (4) Percentage, Graphical, and Statistical Inference, where percents, graphs, and an alpha level at which significance was tested were presented.

Instrumentation

After reading the program description and recommendations, participants were asked to complete an attached reaction sheet measuring the extent of their agreement with the recommendations, their reactions to the evaluation's usefulness, information utilization, and their ratings of the evaluation presentation. All items were Likert type having "Strongly Agree" rated as a "1" and "Strongly Disagree" rated as a "5". A total score was calculated for each of the dependent measures by summing the items composing each dependent construct. Variations of this instrument has consistently had an alpha reliability coefficient above .85. The alpha reliability coefficient for this analysis was .92. Logical construct validity has been shown in previous studies (Brown & Newman, 1982).

RESULTS

A three factor 2x2x4 Multivariate Analysis of Variance was used to analyze the four dependent measures: Extent of Agreement with Recommendations, Ratings of the Perception of the Information Provided, the Utilization of Information, and the Ratings of the Presentation. The fixed factors were: Level of Data, (No Data; Percentage Data; Percentage and Graphical

Data; Percentage, Graphical, and Inferential Statistics), Model Used in the Presentation of the Recommendations (Advocacy; Adversary), and Participant Occupation (Administrator; Teacher). All assumptions for a multivariate analysis of variance were met. This included randomization of subjects, normalacy of the dependent variable, independence between groups. Examination of the error correlation matrix, as suggested by Finn and Mattsson (1978), indicated that a multivariate construct was formed, thus multivariate techniques were used. Table I presents the means and standard deviations. Table II provides the results of the Multivariate Analysis of Variance.

An examination of the three way interaction ($F=1.74$; $df=12,801.95$; $p > .05$) and Data by Occupation ($F=1.01$; $df=12,801.95$ $p > .05$) indicated nonsignificance. The interactions of Model by Occupation ($F=2.69$; $df=4,303$; $p < .05$) and Model by Data ($F=1.82$; $df=12,801.95$; $p < .05$) were found to be significant.

As suggested by Finn and Mattsson (1978), in order to investigate the effects of the interaction of Model by Occupation on individual dependent variables, post hoc univariate analyses were performed. The main contribution to differences among the interaction

of model and subject's occupation was made by the dependent variable of Extent of Agreement with Recommendations ($F=4.99$; $df=1,306$; $p < .05$). Administrators receiving the advocacy wording were in more agreement with the recommendations ($\bar{X}=8.51$) than were administrators receiving the adversary wording ($\bar{X}=10.43$), or teachers receiving either the advocacy ($\bar{X}=10.04$) or adversary ($\bar{X}=13.56$) wording. Multivariate eta squared was found to be .03, indicating three percent of the variability in agreement with recommendations can be attributed to the interaction of model and occupation.

For the significant interaction of data and model, recommendation was once again the main source of variance in the dependent construct ($F=2.29$; $df=3,306$; $p > .05$). Those participants receiving the advocacy wording containing the highest level of data (percentages, graphs, and inferential statistics) indicated the greatest agreement with recommendations ($\bar{X}=10.59$) followed by those receiving the advocacy wording with no data ($\bar{X}=10.60$). The lowest agreement came from the subjects receiving the adversary wording with no level of data ($\bar{X}=11.76$) followed by those receiving percentages within the advocacy wording ($\bar{X}=11.63$)

as well as the adversary wording ($X=11.53$). Eta squared was found to be .07.

DISCUSSION AND CONCLUSION

This study investigated the effect of the level of data, evaluation model received, and occupational affiliation on educators' perception of an evaluation report. It should be noted that experimental findings may be limited due to the fact that the vignettes were a "simulated" as opposed to a "real" situation. Previous research (Braskamp, Brown, & Newman, 1982) however, has shown that the use of vignettes under controlled situations is an effective means of studying theoretical propositions about program evaluation. Given these limitations, several conclusions may be drawn.

It appears that it is important for evaluators to be aware of the interaction of factors which influence decision-making. The results of this study showed model to interact significantly with both the level of data and subject occupation. In reviewing the interaction of model by data, the highest agreement toward recommendations appeared when subjects received the advocacy wording using inferential statistics or the advocacy wording receiving no data. The least

amount of agreement came from those subjects receiving the adversary wording with no data. These results suggest that if the recommendation is positive, either the presence of inferential data or no data leads to more agreement. When the recommendations are phrased in a negative manner, however, some form of statistical rationale for the decision is necessary. The most beneficial appears to be the use of percents and graphs. Inferential statistics do not appear to increase agreement.

The interaction of model by participant occupation shows the highest agreement on recommendations to come from administrators receiving the advocacy wording followed by administrators receiving the adversary wording. No major differences were found for teachers when they received the advocacy wording over the adversary wording.

This study supports and expands upon results found in previous research. When studying the effects of the presence of both types of wording, Brown & Newman (1982) found audience agreement toward recommendations to be higher when material was presented in an order using using adversary wording followed by advocacy wording. In the present study where

respondents received only the adversary or advocacy approach but not both, data and occupation were found to interact with the wording in effecting agreement with recommendation information. A second study (Newman & Howell, 1985) showed those educators in general who received the advocacy wording were in more agreement with recommendations made by an evaluator than those receiving the adversary wording. This study found that level of occupation, e.g. administrators verses teachers, interact with model. This supports previous research hypothesizing that the professional level of the audience is an important characteristic in determining evaluation use (Braskamp, Brown, & Newman, 1982; Newman, Bull, Brown, & Rivers, 1984; Newman, Brown, & Littman, 1979).

An even closer comparison can be made with Brooks' (1984) finding that occupation interacted with model to effect agreement toward recommendations. In that study, principals and teachers were more in agreement toward recommendations when presented in an advocacy approach than were administrators and evaluators.

The implications of this study effect practicing evaluators as well as research on evaluation. The results of this study, as well as previous studies,

need to be considered before programs are to be evaluated. For the evaluator, model type as well as level of data should be of concern in deciding how the evaluation should be presented to varying audiences. It appears that, for administrators, evaluations using the advocacy wording will gain higher agreement with recommendations while teachers will not be influenced by the wording. Overall, the use of statistics to support recommendations appears to be of greater importance when using the adversary wording. For evaluators of vocational education programs, this study implies that more care is needed in determining the contextual variables that may be interacting with utilization. More research needs to be conducted within vocational/technical education which examines other types of programs, e.g. curricular or administrative, other audiences, and other models of evaluation. Other utilization variables that could be studied include funding, retro-active decision-making and increased uses of evaluator skills.

This study reconfirms the fact that good program evaluation and perceptions of program evaluation are dependent on many variables. The type of report, the design of evaluation, and audience characteristics all merge to make program evaluation a multi-faceted

activity. Future research should focus not only on these variables, but on others that may cause differences in perception of evaluation reports to appear.

REFERENCES

- Alkin, M. C., Daillak, R., & White, P. (1980). Using evaluations: Does evaluation make a difference? Beverly Hills: Sage.
- Braskamp, L., Brown, R., & Newman, D. L. (1978). Credibility of local educational program evaluation report: Author source and client audience characteristics. American Educational Research Journal, 15, 441-450.
- Brooks, C. R., (April 1984). Administrative decision-making and program evaluation use. Paper presented at the American Educational Research Association Meeting. New Orleans.
- Brown, R., Braskamp, L., & Newman, D. L. (1978). Evaluator credibility as a function of report style: Do jargon and data make a difference. Evaluation Quarterly, 2, 331-341.
- Brown, R., & Newman, D. L. (1982). An investigation of the effect of different data presentation formats and order of arguments in a simulated adversary evaluation. Educational Evaluation and Policy Analysis, 2, 197-203.

- Carter, R. K. (1971). Clients' resistance to negative findings and the latent conservative function of evaluation studies. The American Sociologist, 6, 118-124.
- Finn, J. D., & Mattsson, I. (1978). Multivariate analysis in educational research: Applications of the multivariate program. Chicago: National Educational Resources.
- Fishbein, M. & Ajzen, I. (1975). Beliefs, attitudes, invention and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley.
- Hawkins, J. D., Roffman, R. A., & Osborne, P. (1978). Decision-makers' judgements: The influence of role evaluative criteria and information access. Evaluation Quarterly, 2, 435-455.
- House, E. R. (1978). Assumptions underlying evaluation models. Educational Researcher, 7, 4-11.
- Kourilsky, M. L. (1973). An adversary model for educational evaluation. Evaluation Comment, 1, 3-6.
- McGuire, W. J. (1969). The nature of attitude change. In G. Lindzey and E. Aronson (Eds.), The handbook of psychology. Reading, MA: Addison-Wesley.
- Newman, D.L., Brown, R., & Littman, M.I. (1979). Evaluator, report, and audience characteristics

which influence the impact of evaluation reports:

Does who says what to whom make a difference?

CEDR Quarterly, 12, 14-18.

Newman, D. L., & Howell, B. E. (1985). The effect of wording recommendations and use of data on perceptions of evaluation: An empirical study. Paper presented at the National Evaluation Conference. Toronto.

Rutman, L. (1982). Dimensions of utilization and types of evaluation approaches. Studies in Educational Evaluation, 8, 163-168.

Thompson, B. & King, J. A. (April 1981). Evaluator types: Do differences matter? Paper presented at the Annual Meeting of the American Educational Research Association. Los Angeles, CA.

Wolf, R. L. (1979). The use of judicial evaluation methods in the formation of educational policy. Educational Evaluation and Policy Analysis, 1, 19-29.

Worthen, B. R. & Sanders, J. R. (1973). Educational evaluation theory and practice. Worthington, OH: C. A. Jones.

TABLE I
MEANS AND STANDARD DEVIATIONS

	Recommendations			Evaluation's Usefulness			Utilization of Information			Presentation Ratings		
	n	\bar{X}	S	n	\bar{X}	S	n	\bar{X}	S	n	\bar{X}	S
No Data												
Advocacy												
Administrator	19	7.89	3.11	19	9.53	2.67	19	10.11	3.84	19	8.21	2.22
Teacher	19	9.15	3.17	19	11.68	2.56	19	12.73	3.36	19	10.84	2.22
Adversary												
Administrator	26	11.27	3.55	26	10.50	4.13	26	11.65	4.06	26	9.38	3.12
Teacher	24	13.13	3.42	24	13.13	3.78	24	13.50	4.21	24	11.88	2.64
Percents												
Advocacy												
Administrator	15	9.33	4.19	15	10.53	4.10	15	12.46	4.32	15	9.47	3.78
Teacher	26	9.77	3.56	26	12.27	3.78	26	12.54	3.99	26	11.15	3.50
Adversary												
Administrator	16	8.63	2.94	16	10.19	2.88	16	10.25	2.81	16	8.63	2.39
Teacher	14	12.29	1.33	14	13.07	1.90	14	13.43	2.85	14	11.71	2.73
Percents & Graphs												
Advocacy												
Administrator	20	8.55	4.77	20	10.45	3.52	20	9.30	4.04	20	8.90	3.77
Teacher	19	10.31	2.81	19	11.84	2.79	19	11.47	3.01	19	11.26	2.38
Adversary												
Administrator	18	10.06	1.98	18	10.11	3.25	18	11.89	3.07	18	10.11	3.76
Teacher	19	14.16	3.32	19	12.32	3.42	19	12.95	3.55	19	11.42	3.29
Percent, Graph, Inferential												
Advocacy												
Administrator	20	8.45	3.33	20	8.75	2.81	20	9.30	4.01	20	8.40	3.36
Teacher	19	11.00	2.60	19	12.53	2.95	19	13.37	3.34	19	11.53	2.89
Adversary												
Administrator	21	11.09	3.63	21	11.71	2.41	21	11.90	3.36	21	11.33	3.14
Teacher	27	14.19	3.09	27	11.00	3.71	27	12.22	3.75	27	10.19	3.86

1=Strongly Agree 5=Strongly Disagree

TABLE II
 TEST OF MODEL, LEVEL OF DATA, OCCUPATION
 AND THEIR INTERACTIONS

Independent Variable	Dependent Variable	Multivariate F	Univariate F	Degrees of Freedom
Data	Recommendation	1.25	1.83	3/306
	Information		.31	3/306
	Vote		.63	3/306
	Evaluation		.19	3/306
Model	Recommendation	11.82	46.05*	1/306
	Information		2.22	1/306
	Vote		3.85	1/306
	Evaluation		2.94	1/306
Occupation	Recommendation	13.01	39.07*	1/306
	Information		28.97*	1/306
	Vote		21.40*	1/306
	Evaluation		29.72*	1/306
Model by Occupation	Recommendation	2.69*	4.99*	1/306
	Information		.48	1/306
	Vote		.59	1/306
	Evaluation		2.03	1/306
Data by Occupation	Recommendation	1.01	.80	3/306
	Information		.32	3/306
	Vote		.17	3/306
	Evaluation		1.05	3/306
Data by Model	Recommendation	1.82*	2.29	3/306
	Information		.49	3/306
	Vote		1.67	3/306
	Evaluation		.52	3/306
Data by Model by Occupation	Recommendation	1.74	.77	3/306
	Information		3.36*	3/306
	Vote		2.76*	3/306
	Evaluation		2.86*	3/306

*p < .05

CHAPTER 2

EXTENDED REVIEW OF LITERATURE

Introduction

Evaluation is an important aspect of education. Many decisions are made concerning the implementation of programs by use of evaluation techniques (Rutman, 1982). This utilization of evaluation may vary according to the approach used for evaluation and its purpose, the communication model implemented, and the particular program and audience present. It would be useful if evaluators knew which form of communication would be most beneficial in helping the audience with the decision-making process (Alkin, 1980). One way of examining communication is through the use of the communication paradigm (Fishbein & Ajzen, 1975). This model suggests that "who says what to whom" has an effect on how information is received. The message source, the medium by which it is transmitted, and to whom it is being transmitted, all play an important role in influencing the decisions to be made (McGuire, 1969). Hawkins, Roffman, and Osborne (1978), have

found different media and modes to be useful in creating an effective communication model.

The major purpose of searching for an effective communication model in evaluation is to aid in the presentation of information and feedback. House (1973) stated that administrators and teachers are not influenced by the feedback they receive from evaluators; instead teachers and administrators feel evaluations are quick to point out things which are not being accomplished in the program but fail to point out the accomplishments of the program. In other cases, it has been theorized that evaluation is not properly utilized because of the slow process of obtaining evaluation results (Davis & Salasin, 1975).

Evaluation utilization is not only dependent upon the communication model and its accompanying procedure, but also upon audience variance. Acceptance of the evaluation by the audience is affected by both the audiences' characteristics and self interests. Russell (1981) found the need for differing information formats according to the audiences' job title. For example, administrators were interested in school district performance and summation of past events while teachers were interested in what would affect them personally and in the future.

Statement of the Problem

The purpose of this study was to examine the relationship between the type of data presented, the type of evaluation method used, and specific characteristics of the audience on utilization of evaluation. A hierarchy of data received was varied among administrators and teachers within two structures of the judicial evaluation model, that of Advocacy and Adversary. The construct of concern in studying these variables was their effect on: 1) respondents' extent of agreement with the evaluators recommendations; 2) respondents' perception of the clarity of information provided; 3) respondents' utilization of information; and 4) respondents' perception of the overall presentation of the evaluation information.

Review of Literature

For a program to be evaluated, some form of evaluation method must be applied. Datta (1981) indicated a need for research to be conducted to determine variables which may have an effect on the decision-making process. Newman, Brown, and Rivers (1980), indicate that some variables in need of further study are: the way in

which reports are presented to the audience, the type of information used in the presentation, and the audience for whom the report is being prepared. Following is an indepth look at each of these variables.

Models in Evaluation

Research has shown that different evaluation models do indeed produce different outcomes, and should be prepared and presented differently (Worthen & Sanders, 1973). While a plethora of models, theories and designs of evaluation have been developed, there are three models frequently used in education: Scriven's formative-summative evaluation, Tyler's rationale model, and the judicial evaluation method (Posavac & Carey, 1980). All of these models are used to help the audience in making decisions concerning the program under evaluation.

To establish and justify the merit and worth of a program, Scriven (1967) developed the formative and summative approaches of evaluation. In the formative approach, decisions are made when implementing a program; evaluators use a feedback approach to improve the program while in process. In the summative approach, decisions are made concerning whether a program should

be started, continued, or which program among many would be best to implement. This model uses information which is utilized after the fact, i.e. summarizing the project (Scriven, 1970).

The Tyler rationale was formulated to compare student performance with behaviorally stated objectives. The major goal was to determine the extent to which the purposes of a learning activity were actually being realized. In this model, the evaluator is a curriculum specialist frequently using a pre-post measure of performance.

In this particular study two models which vary by type of recommendation will be utilized: the advocacy model and the adversary model. These models are components of a larger model, the judicial evaluation method which was first developed for federal hearings of national programs. It is based on the legal/judicial mode of decision-making and involves the equivalent of a judge, jury, and opposing opinions in the interpretation of data and program results. In the advocacy approach, proposed by Wolf (1979), the evaluator presents the material in a positive or 'pro' manner. At the opposite end of the judicial spectrum lies the adversary approach in which material is presented

to the audience in a negative or 'con' manner (Kourilsky, 1973). The use of the judicial model has been advocated and implemented by many field agencies (Levine, 1976; Worthen & Owens, 1978; Popham & Carlson, 1977; Thurston, 1978). Brown and Newman (1982), found that decision-making in an evaluation context can be influenced by the order in which the arguments (pro-con as compared to con-pro) are presented in an evaluation. When the adversary argument was presented followed by the advocacy argument (con-pro), audience support was more in favor of recommendations made toward an inservice training program.

Data Techniques

The amount and degree of sophistication of data to be used and/or presented in evaluative settings has been an area of concern for evaluators. In a study by Brown, Braskamp, and Newman (1978), it was found that the use of data interacted with the language used in educators' perception of the evaluation. Educators included both teachers and public school administrators. Language was defined as being either jargon-loaded or jargon-free. Jargon-loaded reports were defined as those that used professionally derived terms to

succinctly convey a concept to a particular group, whereas jargon-free reports used lay language. These styles of language were crossed with variations in empirical presentations. Objective reports were defined as those reporting percentages with findings. They included statements such as "Seventy-five percent of the teachers favored..." and "Among the parents responding to a national survey, 35 percent opposed..." Subjective reports presented no percentages and were more evaluator opinionated. Subjective statements included phrases such as "I believe," "I think," and "In my opinion." The results of this study showed that jargon-loaded subjective reports were rated by the audience to be more technical than jargon-loaded objective, and were found to be the most difficult type of report to read. The report rated as easiest for the audience to read and understand was the jargon-free subjective report. Thus, type of information did affect audience perception of the evaluation; however, it did not influence acceptance of evaluator's recommendations. In addition, the audience did not vary in ratings of the evaluator no matter what report style was used.

In a second study, utilizing levels of data beyond subjective and objective, four different variations in data styles were examined and found to influence decision-making in an evaluation context. Data presentations consisted of four levels: (1) No Data; (2) Percentage Data; (3) Percentage and Graphical Data; and (4) Percentage, Graphical, and Statistical Inference. Results showed that the audience agreed more with the evaluator's recommendations when the data presented were in the form of percentages and graphs. When an inferential statement was included, the level of agreement was lower than any other style of data usage, even that of the no data report (Brown & Newman, 1982). There were, however, no significant main effects on the dependent measure of the usefulness of information in this study.

Brooks (1984) also hypothesized that varying data presentation would affect educators' responses to recommendations. Again, data presentation was varied by four levels of data: (1) No Data, (2) Percentage Data, (3) Percentage and Graphical Data and (4) Percentage, Graphical, and Statistical Inference. This study found no differences in response to the recommendations and ratings as a result of variations in data presentations.

Alkin and Stecher (1983) have hypothesized a decision-making cycle containing a recognition phase, decision-making phase, and ratification phase. They theorize that within the decision-making phase evaluation information is seldom used. In this stage the factors most highly relied upon were beliefs and opinions (i.e. subjective data) which excluded such data information as percentages or graphs.

Audience Needs in Program Evaluation

Several authors (Maudaus, Scriven, & Stufflebeam, 1983; Alkin, 1980), have indicated that the role of the evaluation audience is crucial in educational decision-making. For example, Tyler's major audiences are managers and psychologists, Stufflebeam's audiences are decision-makers, (Maudaus, Scriven, & Stufflebeam, 1983), while the major audience of Wolfs' judicial evaluation method consist of a jury (Wolf, 1979). Several different definitions of audience in the evaluation process are available (House, 1978). Despite these differences in definitions, all theorists indicate that audience perceptions of evaluation and its usefulness are key elements in improving utilization (Maudaus, Scriven, & Stufflebeam, 1983). A study by Brown,

Newman, and Rivers (1980), has shown that when an audience was provided with information about the importance of evaluation, it was more satisfied with evaluative information, and also rated the evaluator higher than did an audience reading the same report without this information. The study also found that people who feel they do not have control over a decision-making process, that is are of external control, are less apt to consider evaluators' recommendations toward decisions. On the other hand, those who feel they are in control will support evaluator recommendations. Title of the information provider has also been shown to affect audience perceptions of evaluation. Audience ratings of the objectivity of the evaluation report are influenced by the description of the person doing the evaluation (i.e. "evaluator", "researcher", or "content specialist"). Teachers and administrators overall rated the "researcher" descriptor to be significantly more objective than the title of "evaluator" or "content specialist" (Braskamp, Brown, & Newman, 1978). In a study by Thompson and King (1981), administrators were found to be more attentive than were teachers to the technical merits contained within a report style than the evaluators' personal characteristics.

Professional level, professional area, and sex are other audience characteristics found to be related to client reactions. In a study by Newman, Brown, and Littman (1979), professional educators were more critical in rating the evaluator than were novice educators. Also, education students rated the evaluation of an education program lower than did business students when the evaluator was a male, but education majors rated the evaluation higher when the evaluator was female. This pattern was reversed for business majors in that business students rated the evaluation of an educational program lower than education students when the evaluator was female, and higher when the evaluator was male. In examining variations in reactions to sex of evaluator by audience role, women evaluators were rated lower by persons whose professional field differed from the content of the evaluation. For example, evaluations conducted by women in the field of business had a lower rating from education readers than those evaluations conducted by men. This supports Carter's (1971) proposition that the closer the audience is to the decision-makers role, the more critical they are of the evaluator, especially of an external evaluator.

Brooks (1984) found a significant difference in both the response to the recommendations and ratings of the evaluator and report as a result of varying organizational roles (teachers, principals, administrators, and evaluators). Teachers, principals, and administrators responded similarly to recommendations made toward a decision while evaluators were in more disagreement with recommendations.

Several studies have also been conducted which examined the relationship of personality characteristics of the audience and evaluation needs and utilization. Newman, Brown, and Rivers (1983) rated decision-makers as either being internally or externally controlled using Rotter's (1966) definition of locus of control. The results indicated that people who perceived themselves as having less control over their decisions in life (i.e. external) did not want as much evaluative information and were less supportive of evaluation in general than did internally controlled individuals who found more information to be of more use. Internally controlled individuals were task oriented, comfortable in making decisions, and listened to others ideas in group work. In a follow-up study, Newman, Bull, Brown, and Rivers (1984) also found internally controlled subjects to

be more supportive of a program than were externally controlled subjects.

Statement of Hypotheses

Based on the above literature, it is hypothesized that the type of evaluation model received will interact with the level of data presented in differentially affecting administrator and teachers' perception of evaluation. The models implemented provide half the subjects with the advocacy wording, while the other half receive the adversary wording. The perceptions of evaluation to be measured in this study concern respondents disagreement toward recommendations of the program, ratings of the information provided, ratings of utilization of information, and ratings of the overall presentation of evaluation information.

CHAPTER 3

METHODOLOGY

Introduction

This chapter presents an extended examination of the research methodology used in the study. It discusses the participants used and describes the instrumentation. A description of the vignette to be used as the treatment is also presented. The chapter concludes with the experimental design and procedures used in collecting and analyzing the data.

Sample

A random sample of 167 teachers and 156 administrators were used for the study. Teachers consists of those currently teaching a vocational education course (i.e. Home Economics, Agriculture, Health, etc.) at the high school level. Administrators were those principals who were located at a high school containing one or more vocational education programs. Each subject randomly received a vignette containing one of the evaluation models (Pro, Con), crossed with a particular

level of data (No Data, Percentage Data, Percentage and Graphical Data, or Percentage, Graphical and Inferential Statistics) included in the simulation.

The suggested sample size, (power at .80, effect size at .40, and alpha at .05) was 18 per cell (Cohen, 1969). There were 16 cells in this analysis necessitating a minimum of 288 participants. Two samples were drawn to achieve the necessary power. The first sample, drawn April, 1985, included a random selection of 200 teachers and 200 administrators. One-hundred-eighteen teachers and 109 administrators returned the vignettes yielding a response rate of 57 percent. A second sample, drawn the following October, included 100 teachers and 100 administrators. Forty-nine teachers and 47 administrators returned the vignettes for a return rate of 48 percent. Sixty-seven percent of the sample analyzed were male while 33 percent were female. Age ranged from 23 to 64 years, the median being 41.

Possible sampling bias may be caused by the distribution process of the vignettes. Vignettes and reaction sheets were mailed to the subjects with the instructions to complete the reaction sheet and return it in an enclosed self-addressed stamped envelope.

Ten days following each vignette mailout, a post card reminder was sent to those participants who had not yet returned the reaction sheet. Participants therefore were involved on a volunteer basis which may affect generalizability to mandatory decision-making settings (Sowell & Casey, 1982).

Vignette

The vignette used in this study was a description of a pilot program for a computer assisted guidance program which had been tried in several schools. The simulated evaluation included information about the teachers' reactions, projected costs of the program, the amount of training required for implementation, and how similar student data were being used before the program was implemented.

Four recommendations made by the program planners were placed within each vignette. These focused on:

1. The number of students a school should have in order to keep the cost of the program at a minimum.
2. Whether or not special teacher inservice programs on computer technology will

make the computerized pupil guidance program more effective.

3. Whether the pupil guidance program should be expanded to include educational and career counseling.
4. If the pupil guidance program should only be implemented in selected grades.

These simulations and recommendations were based on those validated in previous research (Brown & Newman, 1982). Half of the vignettes presented the recommendations using the advocacy (pro) wording and the other half used the adversary (con) wording. In the advocacy approach, the evaluator presented arguments which were in support of the major recommendations, i.e. "All schools with more than 250 students should use the computerized pupil guidance program". In the adversary approach the evaluator presented arguments which were against the major recommendations, for example "Schools must have at least 250 or more students before using the computerized pupil guidance program".

Crossed with each of these models of evaluation (advocacy/adversary) were four different data conditions: (1) No Data; (2) Percentage Data; (3) Percentage and Graphical Data; and (4) Percentage, Graphical,

and Statistical Inference. In the No Data vignette, recommendations were not reinforced with any data which might help in decision-making, e.g. "Few teachers have had experience with computers." Percentage Data sets gave the subjects the percentage number; for example, "Fewer than 30 percent of the teachers surveyed had worked with computers to any extent." Percentage and Graphical Data gave not only the percentage number, but also a graph depicting those percentages. Percentages, Graphical, and Statistical Inference Data contained all of the above with an addition of the alpha level at which significance was tested such as "Statistically, this percentage is significantly less than ($p < .05$) the percentage without computer experience." A sample set of vignettes may be found in Appendix A and B.

Instrument

After reading the program description and recommendations, teachers and administrators were asked to complete an attached reaction sheet measuring the extent of their agreement with the recommendations, their perception of the clarity of information provided, the utilization of information, and how the overall presentation of the evaluation was perceived.

The dependent scales in detail are:

1. Extent of Agreement with the Recommendations.

Four items represented the extent of agreement with recommendations. These included: Whether schools with more than 250 students should use the Computerized Pupil Guidance Program, Whether special teacher inservice training programs on computer technology will make the Pupil Guidance Program more effective, Whether the Computer Pupil Guidance Program should be expanded to include educational and career counseling, and Whether or not a Computerized Pupil Guidance Program should be implemented in all grades within a school.

2. Usefulness of the Information

Provided. Reactions to the Usefulness of the Evaluation Information were measured by four items concerning the description of the program, sufficient information on which to make decisions about the program, enough information to comment on the worth of the program, and whether enough information was provided

so that suggestions could be made on ways to improve the program.

3. The Adequacy of Information. Four items measured whether there was enough information provided so that decisions could be made on the implementation of the program, whether to expand the program, decisions of a budget, and whether or not to provide inservice training for staff.
4. Ratings of the Presentation. The overall ratings of the presentation will be measured by four items in response of the adequacy and fairness of the information received from the evaluator, whether it was adequately technical, was written by a qualified expert, and whether or not it was easy to read and follow.

All items were Likert-type having "Strongly Agree" rated as a "1" and "Strongly Disagree" rated as a "5". Each of the dependent scales were summed to obtain a total score for each subpart. Variations of this instrument has consistently had an alpha reliability coefficient above .85. This particular

study contains an overall alpha reliability of a .92. See Table III for the alpha reliability coefficients for each of the four dependent variables. Logical construct validity has also been shown in previous studies (Brown & Newman, 1982).

Experimental Design

A true experimental posttest-only control group design was used in the study (Campbell & Stanley, 1963). Administrators and teachers were randomly selected from files of the State Department of Vocational and Technical Education within the state of Oklahoma. Random assignment of the type of model and data received were also exercised. Subjects were randomly assigned to groups, exposed to the independent variables, and asked to respond to the reaction sheet. Response scores were then analyzed to determine the effectiveness of the type of model and data received (See Figure 1). Since a random sample of teachers and administrators were drawn throughout the state of Oklahoma, both internal and external validity were controlled. Experimental findings may be limited due to the fact that the treatment was a "simulated", as opposed to "real", situation. Previous research (Braskamp,

TABLE III
ALPHA RELIABILITY COEFFICIENTS FOR THE
FOUR DEPENDENT CONSTRUCTS

Dependent Variable	Alpha Reliability Coefficient
Recommendation	.81
Evaluation's Usefulness	.83
Usefulness of Evaluation	.90
Ratings of the Presentation	.85

<u>HIERARCHY</u>	<u>MODEL</u>	
<u>OF DATA</u>	Advocacy	Adversary
No Data	R X _A O	R X _A O
	R X _T O	R X _T O
Percents	R X _A O	R X _A O
	R X _T O	R X _T O
Graph	R X _A O	R X _A O
	R X _T O	R X _T O
Inferential	R X _A O	R X _A O
	R X _T O	R X _T O

T=Teacher

A=Administrator

R=Random Assignment to Vignettes

X=Vignette

O=Observations

Figure 1. Experimental Design

Brown, & Newman, 1982) however, has shown that the use of vignettes is an effective means of studying theoretical propositions under controlled situations.

Procedure

Each participant was randomly assigned a vignette differing in model of presentation and level of data available for decision-making. The vignette described a decision to be made on a computerized pupil guidance program to be implemented in the classroom. After reading the description of the Computerized Pupil Guidance Program containing recommendations made by an evaluator, the participant completed a reaction sheet using a Likert type format which contained the items relating to the dependent variables. The vignette was mailed to 600 randomly selected administrators and teachers (400 selected in April, 1985; 200 selected the following October). The return rates were 57 percent and 48 percent respectively.

Ten days following the vignette mailout, a reminder was sent to those participants who had not yet returned the reaction sheet, asking that they please do so. A follow-up on the non respondents indicated no major differences for gender or age. In addition, no difference

was found for the April and October samples, thus enabling the two groups to be combined. This resulted in obtaining 118 teachers and 109 administrators in the first sample and 49 teachers and 47 administrators in the second sample, yielding a total sample of 323.

SELECTED BIBLIOGRAPHY

- Alkin, M. C., Daillak, R., & White, P. (1980). Using evaluations: Does evaluation make a difference? Beverly Hills: Sage.
- Alkin, M. C. & Stecher, B. (1983). Evaluation in context information use in elementary school decision-making. Studies in Educational Evaluation, 9, 23-32.
- Braskamp, L., Brown, R., & Newman, D. L. (1978). Credibility of local educational program evaluation report: Author source and client audience characteristics. American Educational Research Journal, 15, 441-450.
- Braskamp, L., Brown, R. & Newman, D. L. (1982). Studying evaluation through simulations. Evaluation Review, 6, 114-126.
- Brooks, C. R., (April 1984). Administrative decision-making and program evaluation use. Paper presented at the American Educational Research Association Meeting. New Orleans.
- Brown, R., Braskamp, L., & Newman, D. L. (1978). Evaluator credibility as a function of report

style: Do jargon and data make a difference.

Evaluation Quarterly, 2, 331-341.

Brown, R., & Newman, D. L. (1982). An investigation of the effect of different data presentation formats and order of arguments in a simulated adversary evaluation. Educational Evaluation and Policy Analysis, 2, 197-203.

Brown, R., Newman, D.L., & Rivers, L. (1980). Perceived need for evaluation as an influence on evaluation's impact on decision-making. Educational Evaluation and Policy Analysis, 2, 67-72.

Campbell, D. T. & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research.
Chicago: Rand McNally.

Carter, R. K. (1971). Clients' resistance to negative findings and the latent conservative function of evaluation studies. The American Sociologist, 6, 118-124.

Cohen, J. (1969). Statistical power analysis for the behavioral sciences. New York: Academic Press.

Datta, L. (1981). Communicating evaluation results for policy decision-making. In R.A. Berk (Ed.), Educational evaluation methodology: The state of the art (pp. 125-145). Baltimore: John Hopkins University Press.

- Davis, J. R., & Salasin, S. E. (1975). The utilization of evaluation. In E. L. Struening and M. Guttentag (Eds.), Handbook of evaluation research, 1, (pp. 621-666). Beverly Hills: Sage.
- Fishbein, M. & Ajzen, I. (1975). Beliefs, attitudes, intention and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley.
- Hawkins, J. D., Roffman, R. A., & Osborne, P. (1978). Decision-makers' judgements: The influence of role evaluative criteria and information access. Evaluation Quarterly, 2, 435-455.
- House, E. R. (1978). Assumptions underlying evaluation models. Educational Researcher, 7, 4-11.
- King, J. A. (1982). Studying the local use of evaluation: A discussion of theoretical issues and an emperical study. Studies in Educational Evaluation, 8, 175-183.
- Kourilsky, M. L. (1973). An adversary model for educational evaluation. Evaluation Comment, 1, 3-6.
- Levine, M. (1974). Scientific method and the adversary model: Some preliminary thoughts. American Psychologist, 29, 661-667.
- Maudus, G. F., Scriven, M., & Stufflebeam, D. L. (1983). Evaluation models: Viewpoints on educational and human services evaluation. Boston: Kluwer-Nijhoff.

- McGuire, W. J. (1969). The nature of attitude change.
In G. Lindzey and E. Aronson (Eds.), The handbook of psychology. Reading, MA: Addison-Wesley.
- Newman, D.L., Brown, R., & Littman, M.I. (1979).
Evaluator, report, and audience characteristics which influence the impact of evaluation reports: Does who says what to whom make a difference?
CEDR Quarterly, 12, 14-18.
- Newman, D. L., Brown, R., & Rivers, L. (1983).
Locus of control and evaluation use: Does sense of control affect information needs and decision-making? Studies in Educational Evaluation, 9, 77-78.
- Newman, D. L., Bull, K. S., Brown, R., & Rivers, L. (April 1984). Locus of control as an influencer of school evaluation needs. Paper presented at the American Educational Research Association.
New Orleans.
- Newman, D. L. & Howell, B. E. (1985). The effects of wording recommendations and use of data on perceptions of evaluation: An empirical study. Paper presented at the National Evaluation Conference. Toronto.

- Popham, W. J. & Carlson, D. (1977). Deep dark deficits of the adversary evaluation model. Educational Researcher, 6, 3-6.
- Posavac, E. J., & Carey, R. G. (1980). Program evaluation: Methods and case studies. Englewood Cliffs, NJ: Prentice-Hall.
- Russell, R. V. (1982). The disparity of information between administrators and teachers. Education, 102, 354-358.
- Rutman, L. (1982). Dimensions of utilization and types of evaluation approaches. Studies in Educational Evaluation, 8, 163-168.
- Scriven, M. (1967). The methodology of evaluation. (American Educational Research Association Monograph No. 110). Lafayette, IN: The Social Science Education Consortium, Purdue University.
- Scriven, M. (April 1980). Educational evaluation and decision-making. Paper presented at the annual meeting of the American Educational Research Association, New York City.
- Sowell, E. J., & Casey, R. J. (1982). Analyzing educational research. CA: Wadsworth.
- Stecher, B. (April 1983). The relationship between evaluative data and school decision-making.

Paper presented at the annual meeting of the American Educational Research Association. Montreal, Quebec.

Thurston, P. (1978). Revitalizing adversary evaluation: Deep dark deficits or muddled mistaken musings. Educational Researcher, 7, 3-8.

Thompson, B. & King, J. A. (April 1981). Evaluator types: Do differences matter? Paper presented at the Annual Meeting of the American Educational Research Association. Los Angeles, CA.

Wolf, R. L. (1979). The use of judicial evaluation methods in the formation of educational policy. Educational Evaluation and Policy Analysis, 1, 19-29.

Worthen, B. R. & Sanders, J. R. (1973). Educational evaluation theory and practice. Worthington, OH: C. A. Jones.

APPENDIX A

ADVOCACY WORDING VIGNETTE

A Computer Pupil Guidance Program
An Evaluation

Description of the Program

Computer Guidance System, Inc. (CGS), a non-profit consortium, has devised a flexible and economical Computer Pupil Guidance Program to assist teachers in making decisions regarding the academic programs of elementary pupils. This program analyzes the pupil's past achievement, expressed career interests, and achievement and aptitude test scores. The Pupil Guidance print-out illustrates pupil strengths which can be further explored, and diagnose pupil weaknesses so remedial help can be provided. The program also prints out recommended courses of action for the teachers and counselors.

The CGS system is flexible, making it possible for any size school to utilize it. The major innovation of CGS is programmed assessments of student learning styles which can be used along with aptitude scores to predict achievement in key academic subjects. Schools with computers may purchase their own taped computer program, whereas others may contract for computer facilities via telephone tie-ins with other state agencies. Each participating school contracts for initial establishment of the Pupil Guidance Program. The school supplies detailed information about its test batteries, grading systems, and record keeping procedures. Curriculum offerings, faculty descriptions, and remedial and special education resources are determined from on-site visits by trained CGS staff. There is a small annual maintenance fee and a standard rate for each student added to the computer file.

Evaluation and Field Testing

A nation-wide board of consultants appointed by CGS worked with the designers of the computer pupil guidance program and persons involved in field testing the program. Evaluative data on the use of the Computer Pupil Guidance Program was gathered from three major school systems (one each in the South, East, and the Midwest). Selected principals and teachers from other individual schools where the Program was field tested were also surveyed and interviewed. Data was analyzed and made available to all potential users.

These vignettes were developed by Robert D. Brown, University of Nebraska, Lincoln, and Dianna L. Newman, Oklahoma State University.

Implementation Decision

The State Commissioner of Education and the State Association of School Superintendents decided to investigate the worth and utility of the Computer Pupil Guidance Program designed by CGS for possible implementation in the state. They agreed that the key evaluation questions were:

1. Should the computerized pupil guidance program be implemented in schools with less than 250 students?
2. Should an inservice training program for teachers which focuses on computers be made a part of the program?
3. Should the computerized pupil guidance program be expanded to include junior high and secondary schools?
4. Should the pupil guidance program be implemented in all grade units within a school?

Representatives from CGS Inc., recommended that the answers to all of the questions be "YES". CGS made all of the evaluation data available to the Commissioner, Superintendents, and their representatives.

Decision-Making Strategy

The State Commissioner of Education and the State Association of School Superintendents decided to employ an evaluation approach referred to as an advocacy model to assist them in answering their questions. In this approach the evaluator presents arguments which support the major recommendations. It is up to the decision-makers to determine whether the arguments are valid or not. Dr. Cockrum, an evaluator who was not affiliated with either the state school or CGS was chosen as the evaluator. Thus, Dr. Cockrum presented arguments which were all in agreement with the recommendations of CGS. Half of the consultant fee was paid by CGS and half by the state.

RECOMMENDATIONS 1-4

ALL SCHOOLS WITH MORE THAN 250 STUDENTS SHOULD USE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

The cost per student will decrease as the number of students in a school system increase and a minimum of 250 students would be sufficient to make the program economical. Slightly larger schools should be able to hire a full-time person to manage the system from the money usually budgeted from the clerical help used to maintain student files.

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL MAKE THE PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

Few teachers have had experience with computers. Because the computer guidance program would reduce teacher clerical work, this saved time could be used for inservice training programs on computer technology.

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

A substantial portion of the school counselor's time is now used in assisting students to understand their test results and help them and their parents make educational and career choices. The computerized program would reduce this time because only students and parents with questions would have to be seen individually.

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED IN ALL GRADES WITHIN A SCHOOL.

Achievement data and test scores are available on students from kindergarten on. Over all grades, the majority of the teachers responding to a survey indicated they would use the system. This was particularly true for teachers in the higher grades.

RECOMMENDATIONS 1-4

ALL SCHOOLS WITH MORE THAN 250 STUDENTS SHOULD USE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

The cost per student decreases with an increase in the number of students and a minimum of 250 students would reduce the extra program expense to less than \$10.00 per student. Estimated savings from money usually budgeted for the clerical help to maintain student files for schools with 500 or more students would be sufficient to pay for a full-time person to manage the guidance system.

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL MAKE THE PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

Fewer than 30% of the teachers surveyed had worked with computers to any extent. It now takes teachers two hours per week to update files and this time could be used for inservice training programs on computer technology.

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

Over half (55%) of the school counselor's time is now used in assisting students to understand test results and help them and their parents to use the data for educational and career choices. The computerized guidance program would reduce this time by 25% because only students and parents with questions would have to be seen individually.

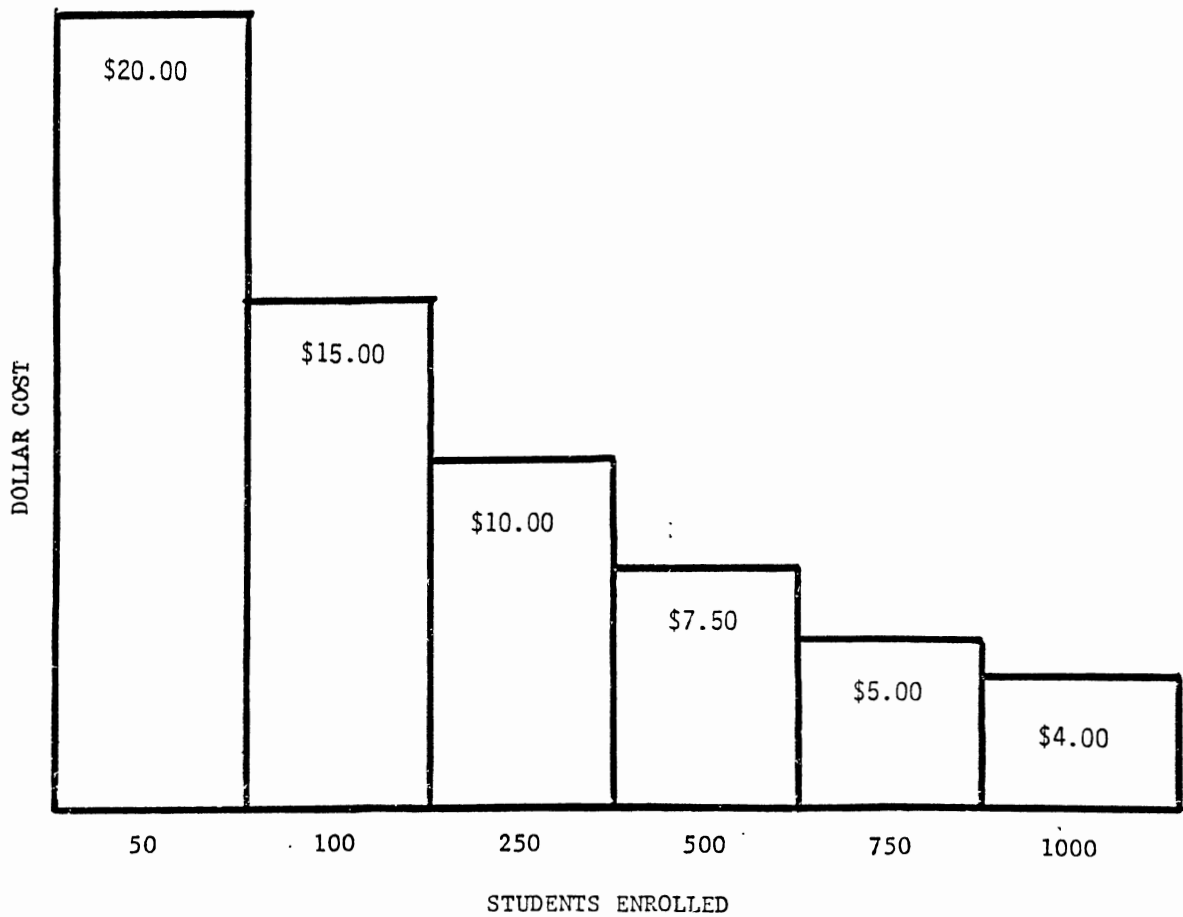
THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED IN ALL GRADES WITHIN A SCHOOL.

Achievement data and test scores are available on students from kindergarten on. Over all grades, the majority (51%) of the teachers responding to a survey indicated they would use the system. This was particularly true for teachers (75%) in the sixth grade.

RECOMMENDATION 1

ALL SCHOOLS WITH MORE THAN 250 STUDENTS SHOULD USE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

The cost per student decreases with an increase in the number of students and a minimum of 250 students would reduce the extra program expense to less than \$10.00 per student (See Figure). Estimated savings from money usually budgeted for the clerical help to maintain student files for schools with 500 or more students would be sufficient to pay for a full-time person to manage the guidance system.

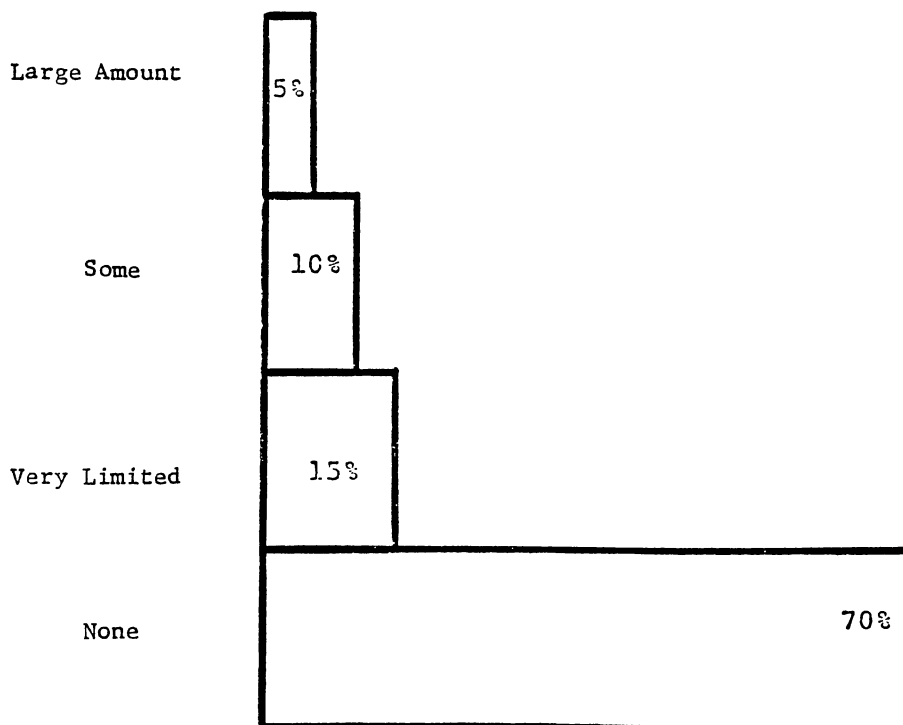


RECOMMENDATION 2

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL MAKE THE PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

Fewer than 30% of the teachers surveyed had worked with computers to any extent (See Figure). It now takes teachers two hours per week to update files and this time could be used for inservice training programs on computer technology.

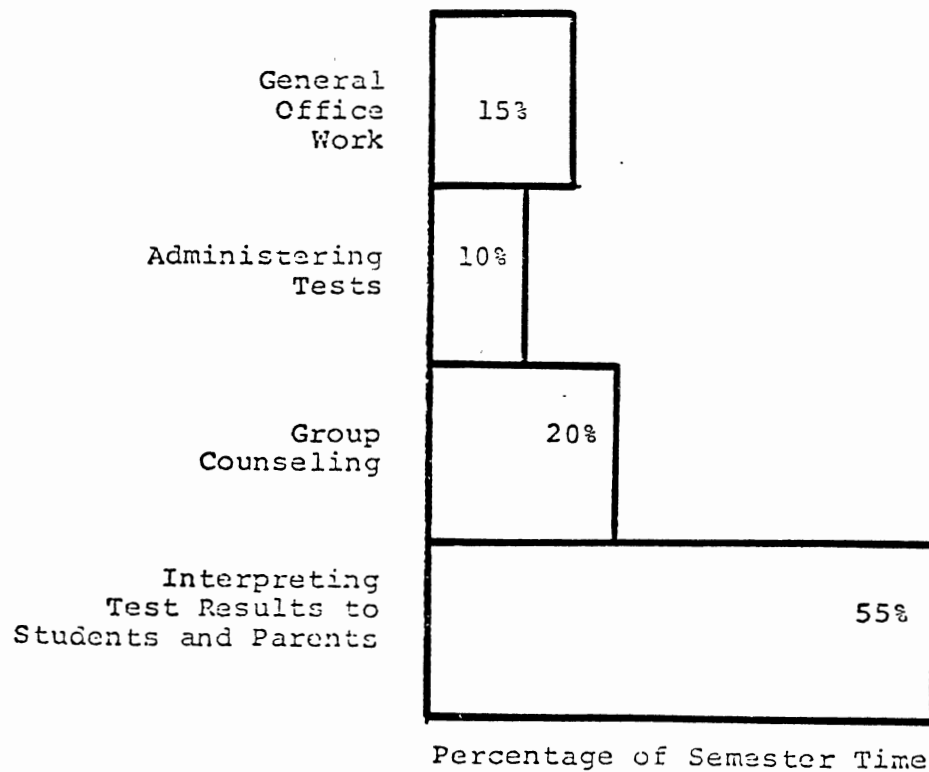
EXTENT OF EXPERIENCE



RECOMMENDATION 3

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

Over half (55%) of the school counselor's time is now used in assisting students to understand test results and help them and their parents to use the data for educational and career choices (See Figure). The computerized guidance program would reduce this time by 25% because only students and parents with questions would have to be seen individually.

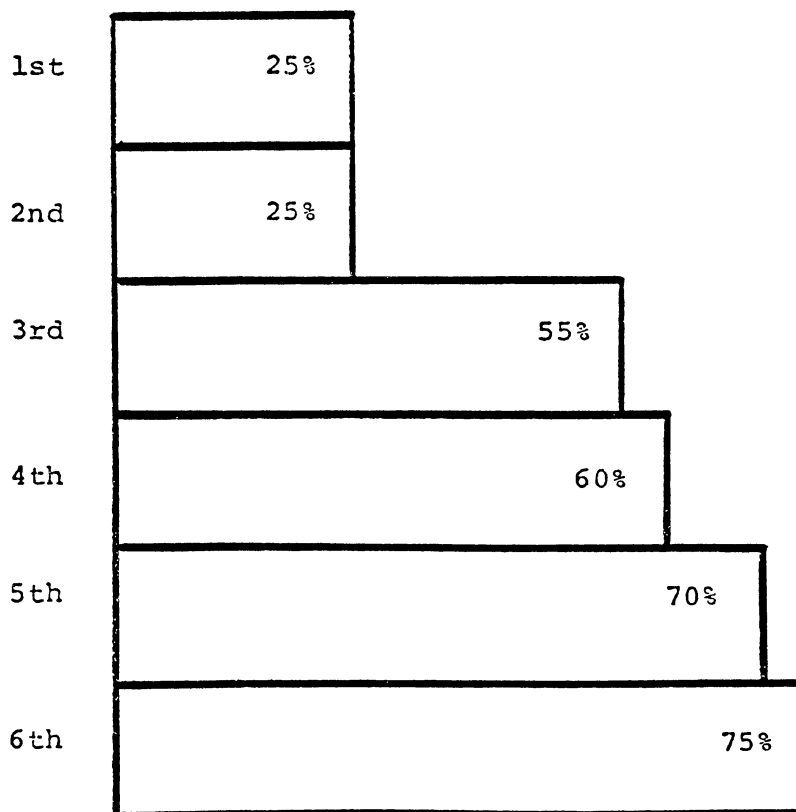


RECOMMENDATION 4

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED IN ALL GRADES WITHIN A SCHOOL.

Achievement data and test scores are available on students from kindergarten on. Over all grades, the majority (51%) of the teachers responding to a survey indicated they would use the system (See Figure). This was particularly true for teachers (75%) in the sixth grade.

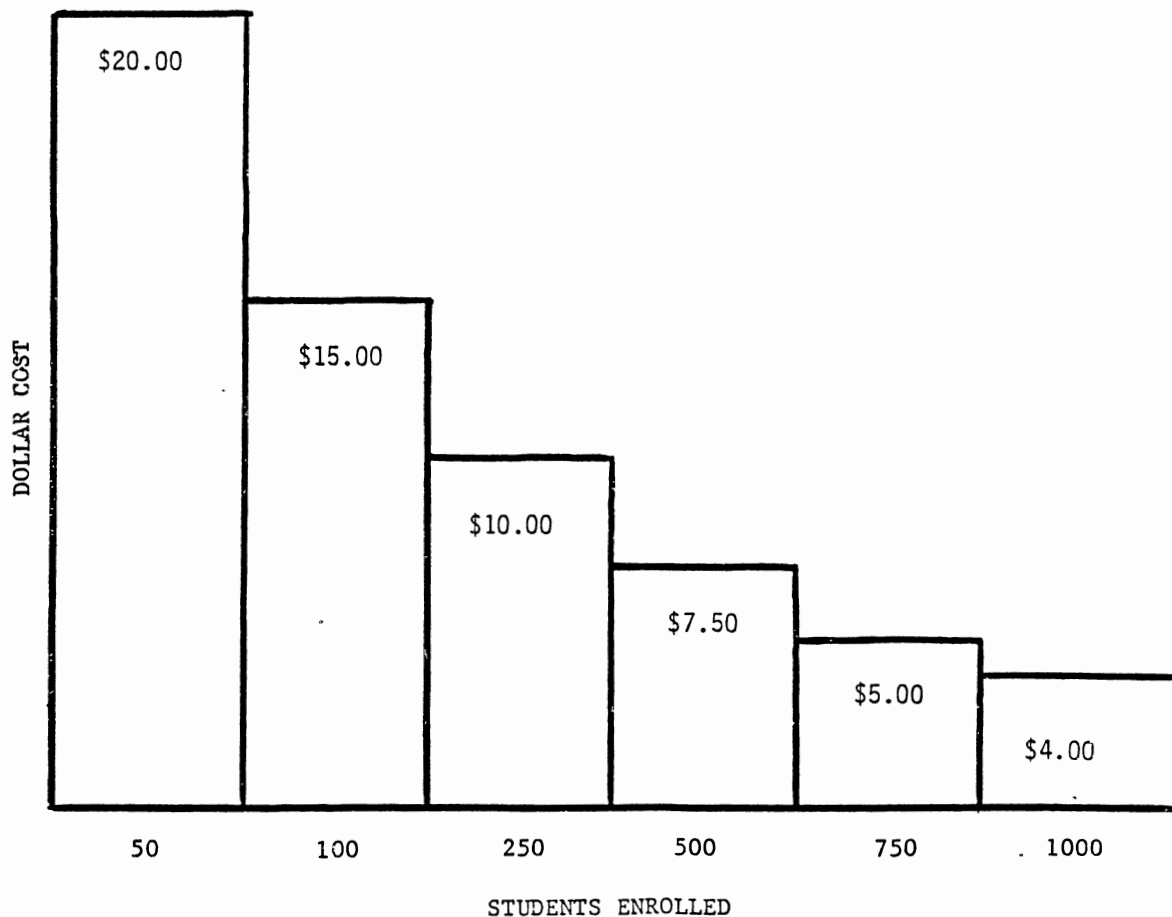
Grade Level



RECOMMENDATION 1

ALL SCHOOLS WITH MORE THAN 250 STUDENTS SHOULD USE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

The cost per student decreases with an increase in the number of students and a minimum of 250 students would reduce the extra program expense to less than \$10.00 per student (See Figure). Statistically, this dollar cost is significantly ($p < .05$) less than the cost for schools with less than 250 students. Estimate. savings from money usually budgeted for the clerical help to maintain student files for schools with 500 or more students would be sufficient to pay for a full-time person to manage the guidance system.

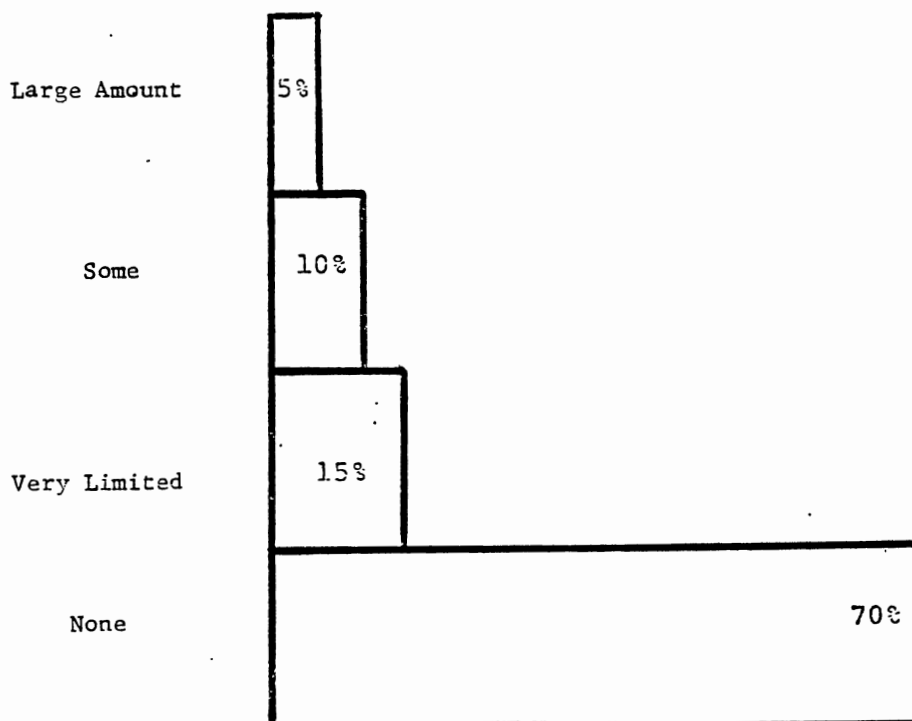


RECOMMENDATION 2

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL MAKE THE PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

Fewer than 30% of the teachers surveyed had worked with computers to any extent (See Figure). Statistically, this percentage is significantly less ($p < .05$) than the percentage without computer experience. It now takes teachers two hours per week to update files and this time could be used for inservice training programs on computer technology.

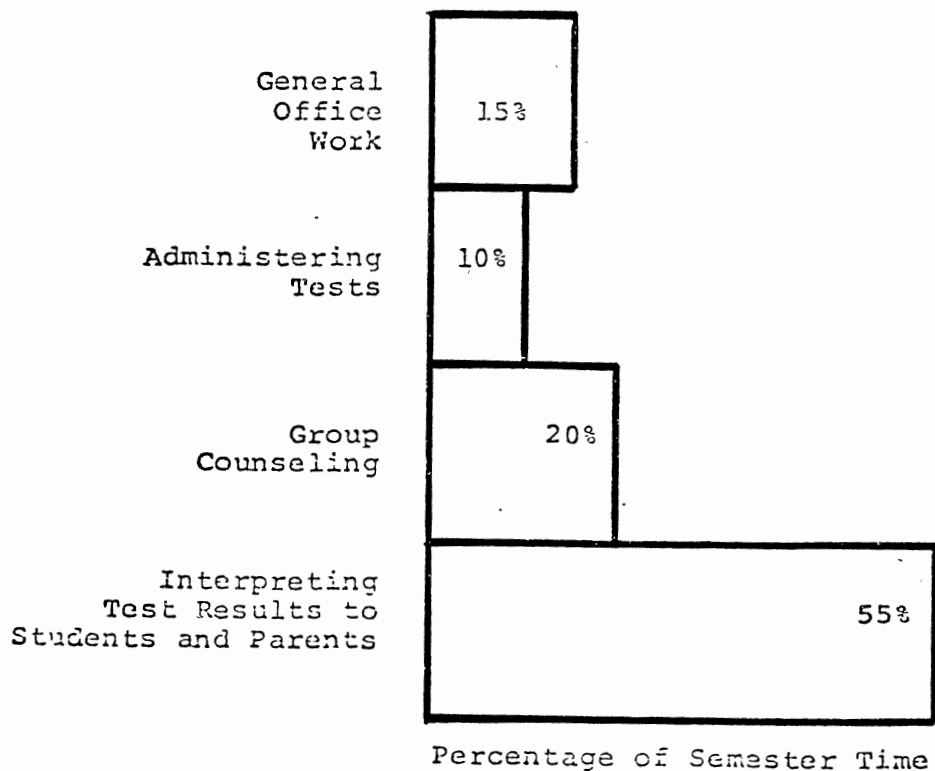
EXTENT OF EXPERIENCE



RECOMMENDATION 3

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

Over half (55%) of the school counselor's time is now used in assisting students to understand test results and helping them and their parents to use the data for educational and career choices (See Figure). This percentage is statistically greater ($p < .05$) than time spent on other testing related activities. The computerized guidance program would reduce this time by 25% because only students and parents with questions would have to be seen individually.

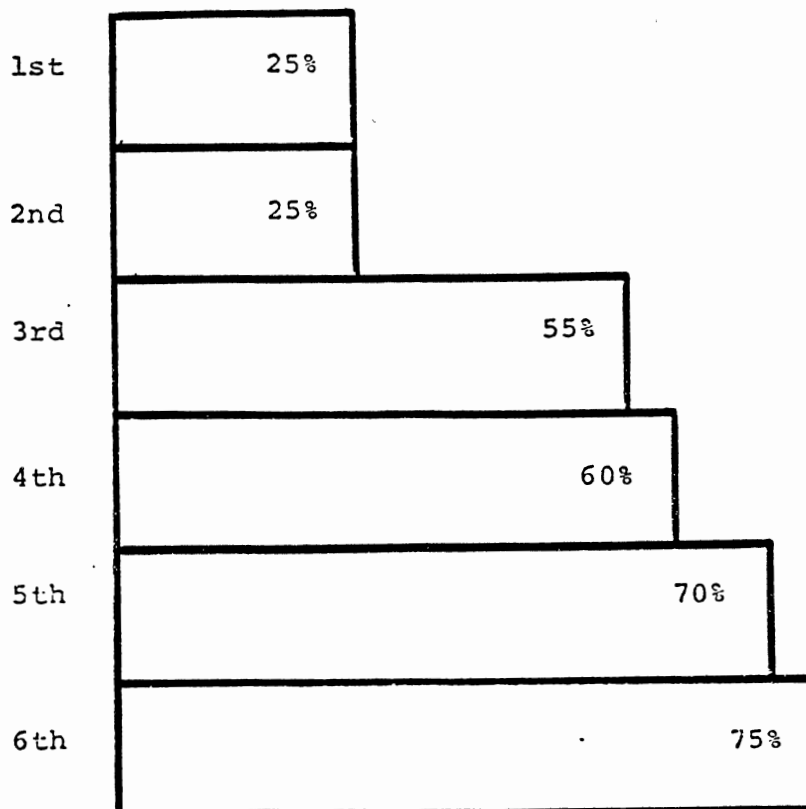


RECOMMENDATION 4

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED IN ALL GRADES WITHIN A SCHOOL.

Achievement data and test scores are available on students from kindergarten on. Over all grades, the majority (51%) of the teachers responding to a survey indicated they would use the system (See Figure). Statistically, this percentage is significantly greater ($p < .05$) than is found among the general population. This was particularly true for teachers (75%) in the sixth grade.

Grade Level



REACTION SHEET

Please give your reactions to the material you read using the following:

SA=Strongly Agree, A=Agree, U=Undecided, D=Disagree, SD=Strongly Disagree

Part I: Extent of Your Agreement with the Recommendations (Circle Your Answer)
(Please refer back to the article)

Recommendation #1	SA	A	U	D	SD
Recommendation #2	SA	A	U	D	SD
Recommendation #3	SA	A	U	D	SD
Recommendation #4	SA	A	U	D	SD

Part II: Reactions to the Evaluation Information
(You may and are encouraged to refer back to the reading material if you wish.)

A. This evaluation information, in my estimation, provided:

1. A clear description of the program.	SA	A	U	D	SD
2. Sufficient information on which to make decisions about the program.	SA	A	U	D	SD
3. Enough so I could comment on the worth of the program.	SA	A	U	D	SD
4. Enough information so I could vote (if a member of the board) on:					
a. whether to implement the program.	SA	A	U	D	SD
b. whether to expand the program to other settings.	SA	A	U	D	SD
c. whether to decide on a budget.	SA	A	U	D	SD
d. whether to provide inservice training for staff.	SA	A	U	D	SD
5. Enough information so I could suggest some ways to improve the program.	SA	A	U	D	SD

B. On the whole I would say that the information was:

1. objective and fair.	SA	A	U	D	SD
2. adequately technical.	SA	A	U	D	SD
3. written by a qualified expert.	SA	A	U	D	SD
4. easy to read and follow.	SA	A	U	D	SD

APPENDIX B

ADVERSARY WORDING VIGNETTE

A Computer Pupil Guidance Program
An Evaluation

Description of the Program

Computer Guidance System, Inc. (CGS), a non-profit consortium, has devised a flexible and economical Computer Pupil Guidance Program to assist teachers in making decisions regarding the academic programs of elementary pupils. This program analyzes the pupil's past achievement, expressed career interests, and achievement and aptitude test scores. The Pupil Guidance print-out illustrates pupil strengths which can be further explored, and diagnose pupil weaknesses so remedial help can be provided. The program also prints out recommended courses of action for the teachers and counselors.

The CGS system is flexible, making it possible for any size school to utilize it. The major innovation of CGS is programmed assessments of student learning styles which can be used along with aptitude scores to predict achievement in key academic subjects. Schools with computers may purchase their own taped computer program, whereas others may contract for computer facilities via telephone tie-ins with other state agencies. Each participating school contracts for initial establishment of the Pupil Guidance Program. The school supplies detailed information about its test batteries, grading systems, and record keeping procedures. Curriculum offerings, faculty descriptions, and remedial and special education resources are determined from on-site visits by trained CGS staff. There is a small annual maintenance fee and a standard rate for each student added to the computer file.

Evaluation and Field Testing

A nation-wide board of consultants appointed by CGS worked with the designers of the computer pupil guidance program and persons involved in field testing the program. Evaluative data on the use of the Computer Pupil Guidance Program was gathered from three major school systems (one each in the South, East, and the Midwest). Selected principals and teachers from other individual schools where the Program was field tested were also surveyed and interviewed. Data was analyzed and made available to all potential users.

Implementation Decision

The State Commissioner of Education and the State Association of School Superintendents decided to investigate the worth and utility of the Computer Pupil Guidance Program designed by CGS for possible implementation in the state. They agreed that the key evaluation questions were:

1. Should the computerized pupil guidance program be implemented in schools with less than 250 students?
2. Should an inservice training program for teachers which focuses on computers be made a part of the program?
3. Should the computerized pupil guidance program be expanded to include junior high and secondary schools?
4. Should the pupil guidance program be implemented in all grade units within a school?

Representatives from CGS Inc., recommended that the answers to all of the questions be "NO". CGS made all of the evaluation data available to the Commissioner, Superintendents, and their representatives.

Decision-Making Strategy

The State Commissioner of Education and the State Association of School Superintendents decided to employ an evaluation approach referred to as an adversary model to assist them in answering their questions. In this approach the evaluator presents arguments which are against the major recommendations. It is up to the decision-makers to determine whether the arguments are valid or not. Dr. Jeffers, an evaluator who was not affiliated with either the state school or CGS was chosen as the evaluator. Thus, Dr. Jeffers presented arguments which were all in disagreement with the recommendations of CGS. Half of the consultant fee was paid by CGS and half by the state.

RECOMMENDATIONS 1-4

SCHOOLS MUST HAVE AT LEAST 750 OR MORE STUDENTS BEFORE USING THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

Although the cost per student will decrease as the number of students in a school system increases, cost effectiveness requires that the minimum number of students would be close to 750 to make the program truly economical. Thus, only large schools will find the program practical.

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL NOT MAKE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

There are enough teachers who have had experience with computers who could assist others, if need be. Special training could be better used for key individuals who would establish and maintain the system.

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD NOT BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

The personal contact counselors have with students and parents as they interpret test and achievement data is important for students and their parents. The amount of time saved by using the computerized system would be limited.

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED ONLY IN SELECTED GRADES WITHIN A SCHOOL.

The program should be implemented only in selected grades. There are a good number of teachers who would prefer not to use it, especially among first and second grade teachers. It should not be forced on teachers who do not want it. Nearly half indicated they would not use it.

RECOMMENDATIONS 1-4

SCHOOLS MUST HAVE AT LEAST 750 OR MORE STUDENTS BEFORE USING THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

Although the cost per student will decrease as the number of students in a school system increases, the cost per student would not be less than \$5.00 until there were over 750 students enrolled. Thus, only schools with 750 students will find the cost reasonable.

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL NOT MAKE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

There are enough teachers (30%) who have had experience with computers who could assist others, if need be. Special training could be better used for key individuals who would establish and maintain the system.

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD NOT BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

The personal contact counselors have with students and parents (55% of their time) as they interpret test and achievement data is important for students and parents. The amount of time saved (25%) by using the computerized system would be limited.

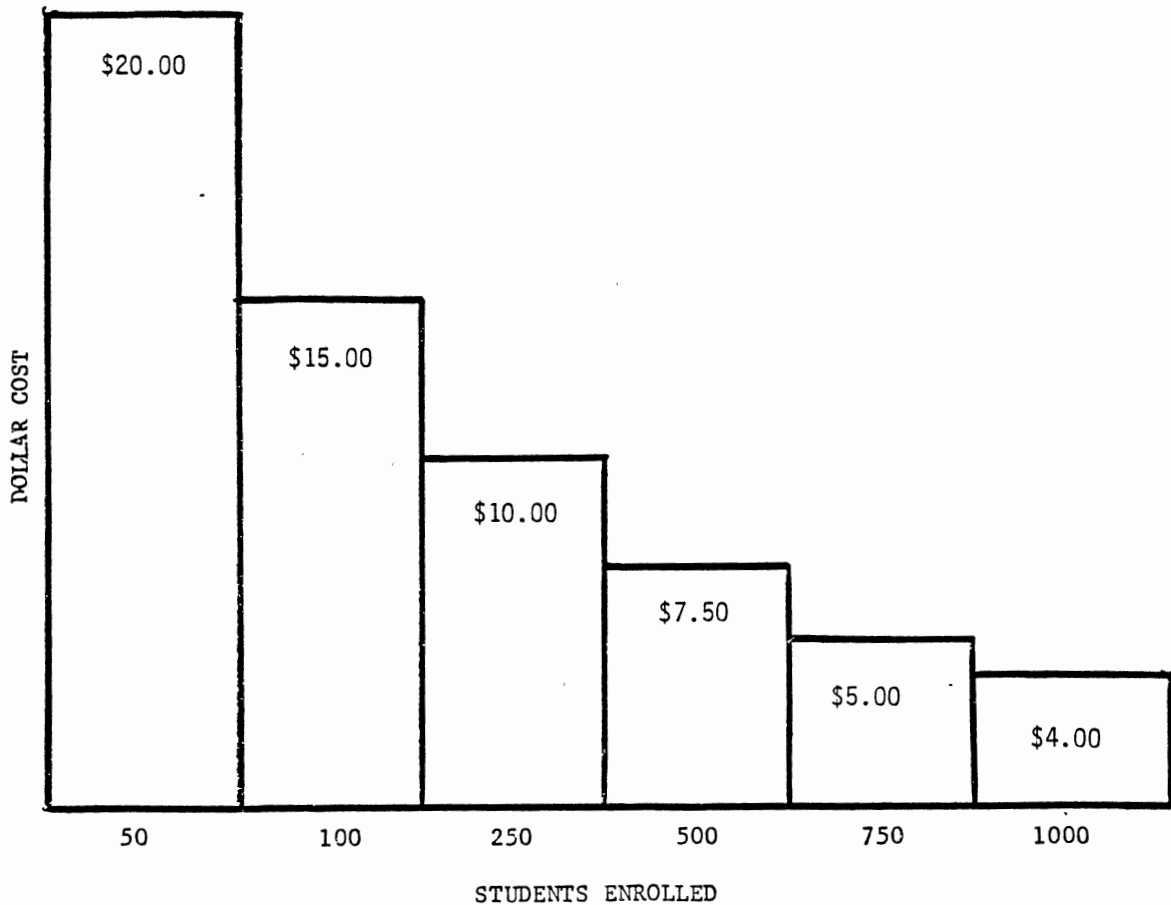
THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED ONLY IN SELECTED GRADES WITHIN A SCHOOL.

The program should be implemented only in selected grades. Only 25% of the first and second grade teachers indicated they would use the computerized program. It should not be forced on these teachers who do not want it. Nearly half (49%) overall indicated they would not use it.

RECOMMENDATION 1

SCHOOLS MUST HAVE AT LEAST 750 OR MORE STUDENTS BEFORE USING THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

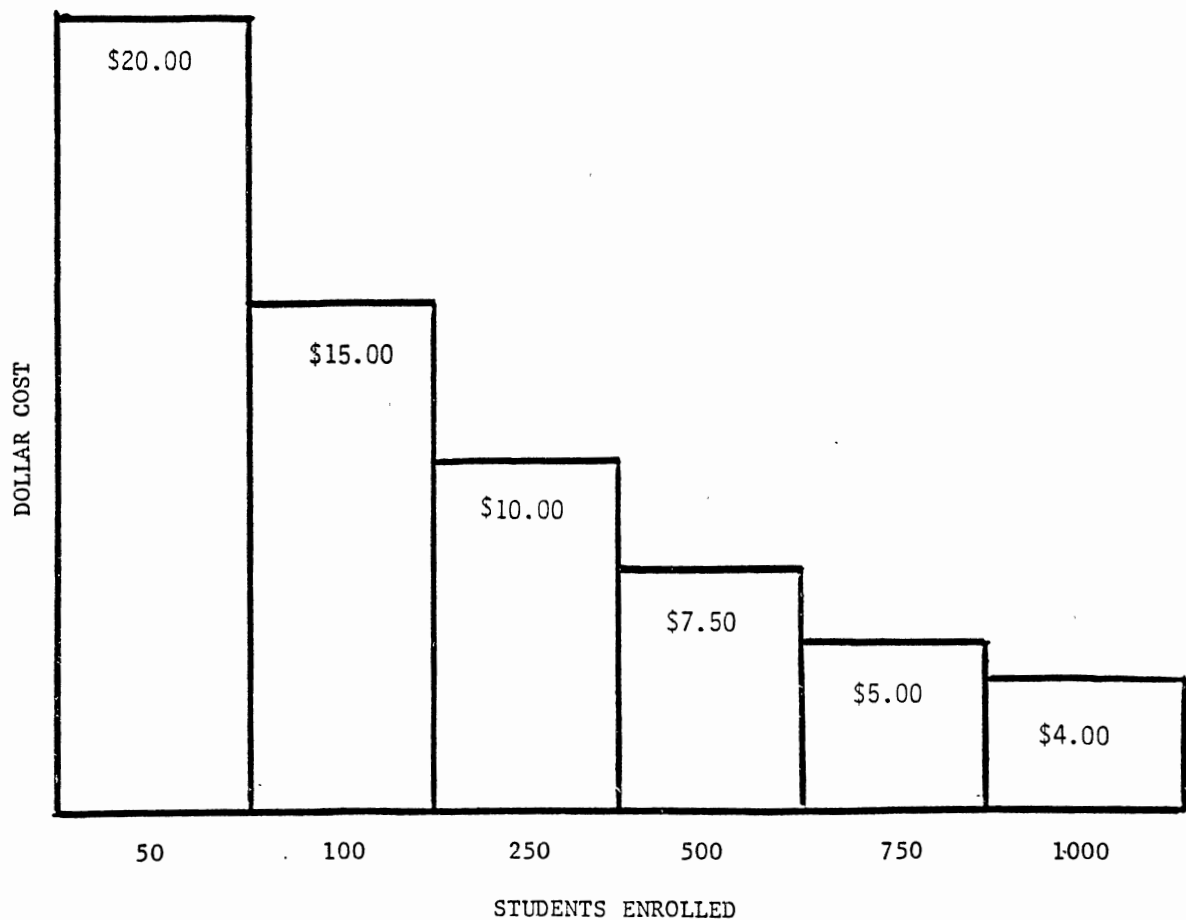
Although the cost per student will decrease as the number of students in a school system increases, the cost per student would not be less than \$5.00 until there were over 750 students enrolled (See Figure). Statistically, this dollar cost is significantly different ($p < .05$) than the cost for schools with less than 750 students. Thus, only schools with 750 students will find the cost reasonable.



RECOMMENDATION 1

SCHOOLS MUST HAVE AT LEAST 750 OR MORE STUDENTS BEFORE USING THE COMPUTERIZED PUPIL GUIDANCE PROGRAM.

Although the cost per student will decrease as the number of students in a school system increases, the cost per student would not be less than \$5.00 until there were over 750 students enrolled (See Figure). Thus, only schools with 750 students will find the cost reasonable.

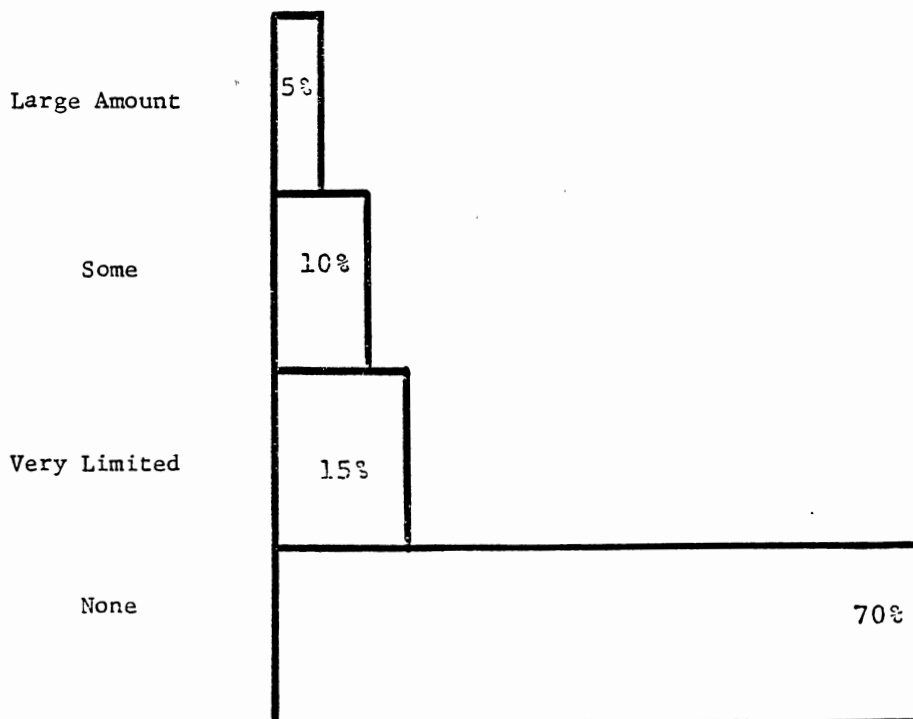


RECOMMENDATION 2

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL NOT MAKE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

There are enough teachers (30%) who have had experience with computers who could assist others, if need be (See Figure). Special training could be better used for key individuals who would establish and maintain the system.

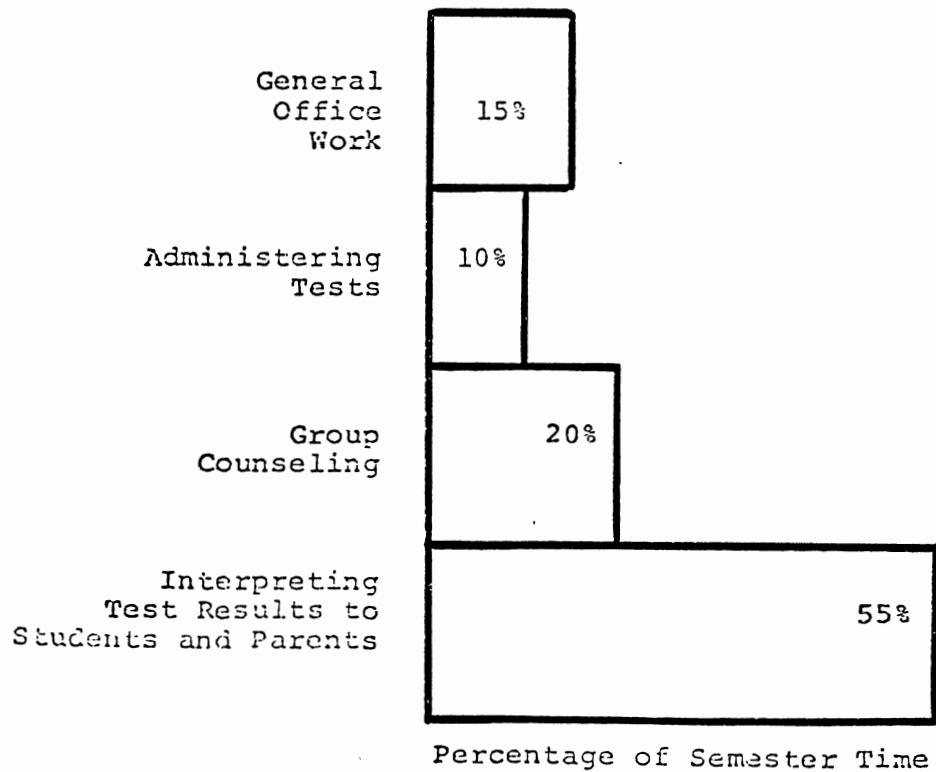
EXTENT OF EXPERIENCE



RECOMMENDATION 3

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD NOT BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

The personal contact counselors have with students and parents (55% of their time) as they interpret test and achievement data is important for students and parents (See Figure). The amount of time saved (25%) by using the computerized system would be limited.

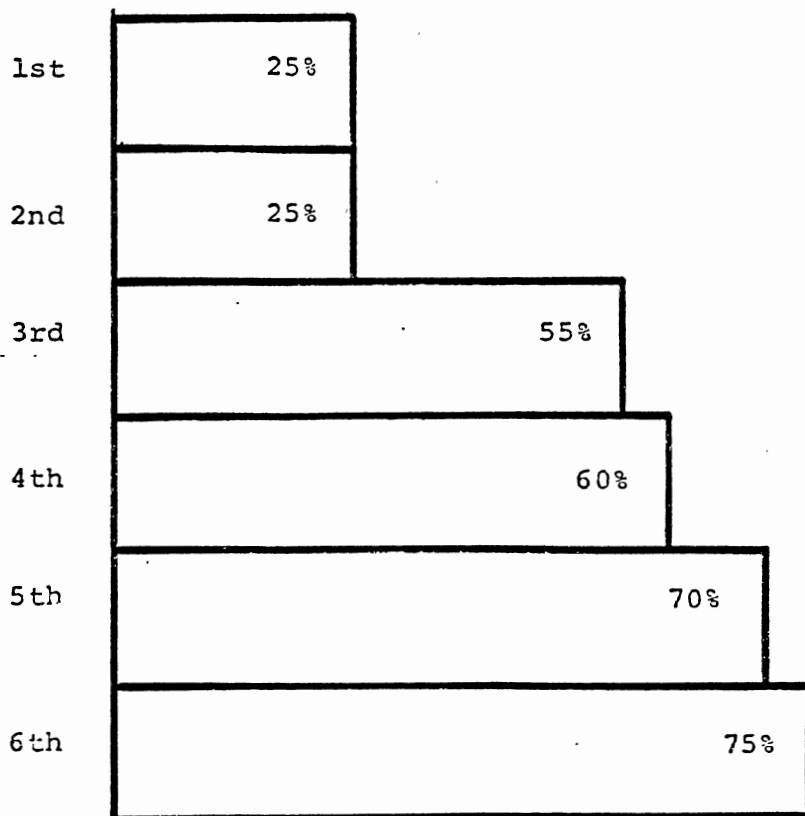


RECOMMENDATION 4

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED ONLY IN SELECTED GRADES WITHIN A SCHOOL.

The program should be implemented only in selected grades. Only 25% of the first and second grade teachers indicated they would use the computerized program (See Figure). It should not be forced on these teachers who do not want it. Nearly half (49%) overall indicated they would not use it.

GRADE LEVEL

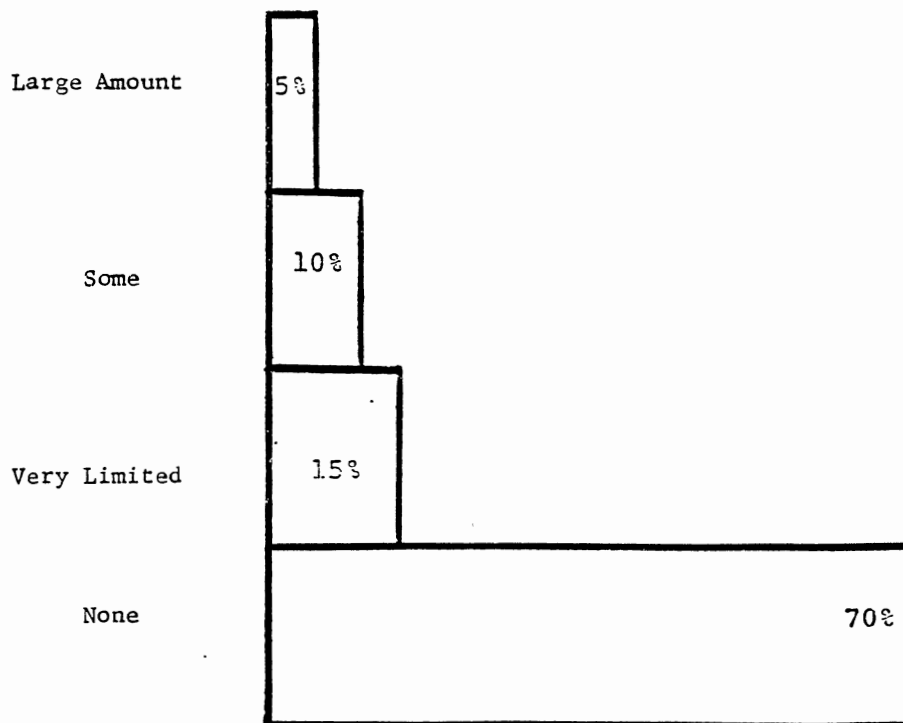


RECOMMENDATION 2

SPECIAL TEACHER INSERVICE TRAINING PROGRAMS ON COMPUTER TECHNOLOGY WILL NOT MAKE THE COMPUTERIZED PUPIL GUIDANCE PROGRAM MORE EFFECTIVE.

There are enough teachers (30%) who have had experience with computers who could assist others, if need by (See Figure). Statistically, this percentage is significantly greater ($p < .05$) than is found among the the general population. Special training could be better used for key individuals who would establish and maintain the system.

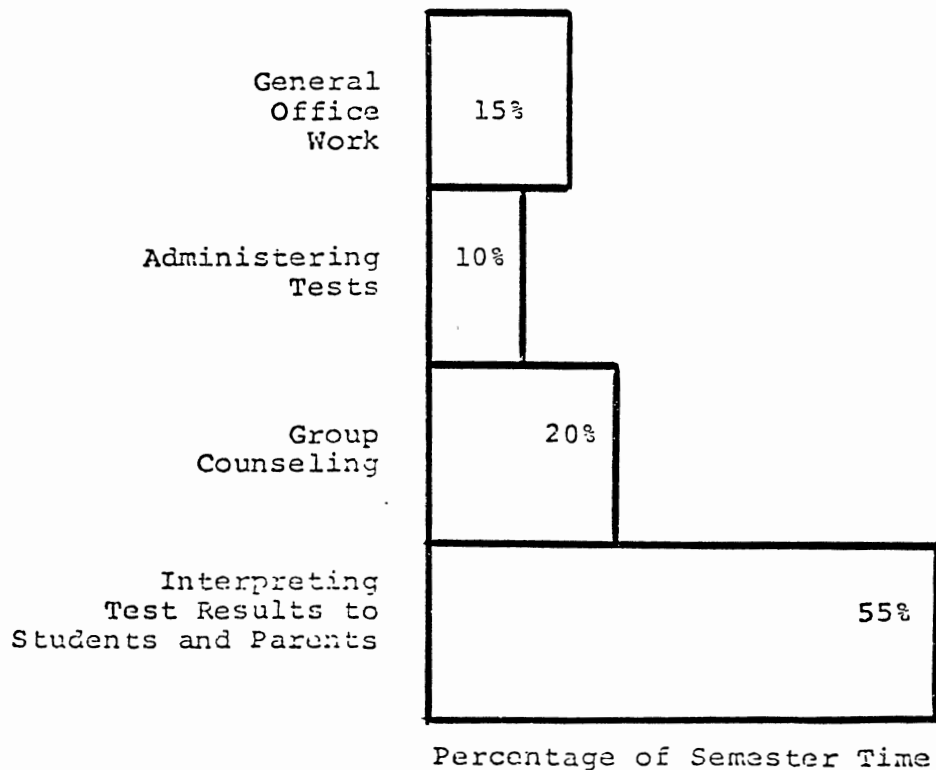
EXTENT OF EXPERIENCE



RECOMMENDATION 3

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD NOT BE EXPANDED TO INCLUDE EDUCATIONAL AND CAREER COUNSELING.

The personal contact counselors have with students and parents (55% of their time) as they interpret test and achievement data is important for students and parents (See Figure). Statistically, this percentage is significantly ($p < .05$) greater than time spent on other testing related activities, which shows how important it is. The amount of time saved (25%) by using the computerized system would be limited.

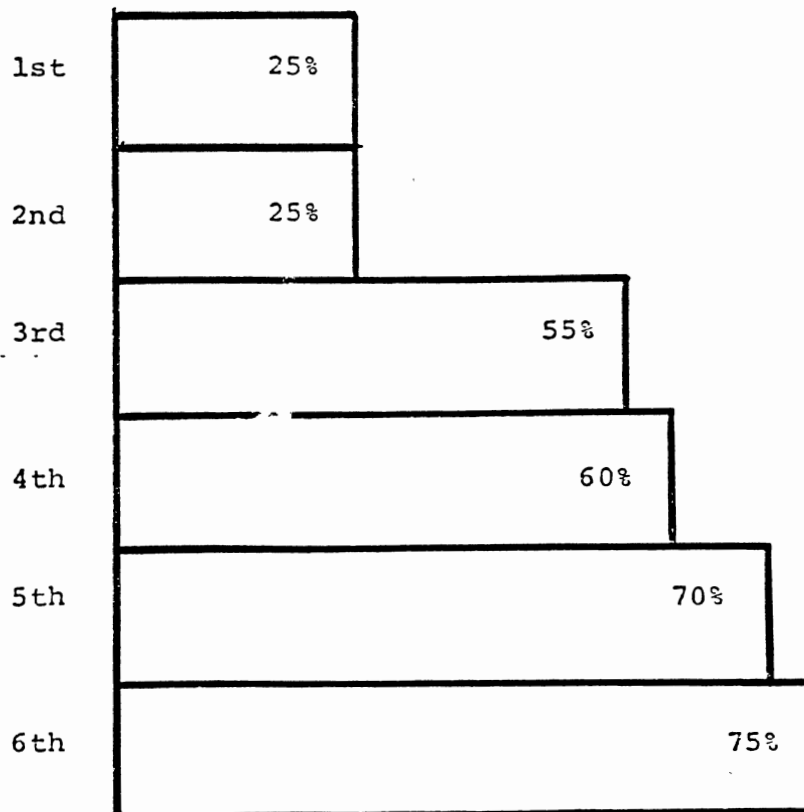


RECOMMENDATION 4

THE COMPUTERIZED PUPIL GUIDANCE PROGRAM SHOULD BE IMPLEMENTED ONLY IN SELECTED GRADES WITHIN A SCHOOL.

The program should be implemented only in selected grades. Only 25% of the first and second grade teachers indicated they would use the computerized program (See Figure). Statistically, this percentage was significantly ($p < .05$) less than for other grades. It should not be forced on these teachers who do not want it. Nearly half (49%) overall indicated they would not use it.

GRADE LEVEL



REACTION SHEET

Please give your reactions to the material you read using the following:

SA=Strongly Agree, A=Agree, U=Undecided, D=Disagree, SD=Strongly Disagree

Part I: Extent of Your Agreement with the Recommendations (Circle Your Answer)
(Please refer back to the article)

Recommendation #1	SA	A	U	D	SD
Recommendation #2	SA	A	U	D	SD
Recommendation #3	SA	A	U	D	SD
Recommendation #4	SA	A	U	D	SD

Part II: Reactions to the Evaluation Information
(You may and are encouraged to refer back to the reading material if you wish.)

A. This evaluation information, in my estimation, provided:

1. A clear description of the program.	SA	A	U	D	SD
2. Sufficient information on which to make decisions about the program.	SA	A	U	D	SD
3. Enough so I could comment on the worth of the program.	SA	A	U	D	SD
4. Enough information so I could vote (if a member of the board) on:					
a. whether to implement the program.	SA	A	U	D	SD
b. whether to expand the program to other settings.	SA	A	U	D	SD
c. whether to decide on a budget.	SA	A	U	D	SD
d. whether to provide inservice training for staff.	SA	A	U	D	SD
5. Enough information so I could suggest some ways to improve the program.	SA	A	U	D	SD

B. On the whole I would say that the information was:

1. objective and fair.	SA	A	U	D	SD
2. adequately technical.	SA	A	U	D	SD
3. written by a qualified expert.	SA	A	U	D	SD
4. easy to read and follow.	SA	A	U	D	SD

VITA

Barbara Ellen Howell

Candidate for the Degree of
Master of Science

Thesis: THE EFFECTS OF EVALUATION MODELS, DATA TECHNIQUES,
AND OCCUPATION ON PERCEPTION AND UTILIZATION
OF EVALUATION TECHNIQUES

Major Field: Applied Behavioral Studies

Biographical:

Personal Data: Born in Oklahoma City, Oklahoma,
January 20, 1961, the daughter of Marvin M. and
Ellen M. Howell.

Education: Graduated from Northwest Classen, Oklahoma
City, Oklahoma, in May, 1978; attended Central
State University, Edmond, Oklahoma, 1978;
attended Oklahoma State University, Stillwater,
Oklahoma, 1979; received Bachelor of Science
in Education degree from Oklahoma State
University, December, 1982; completed requirements
for Master of Science degree at Oklahoma
State University, May, 1986.

Professional Experience: Single Student Housing Desk
Clerk, Oklahoma State University, 1980-1983;
Research and Graduate Assistant, Department of
Applied Behavioral Studies in Education, Oklahoma
State University, 1982-1984; Certified Substitute
Teacher, Stillwater Public Schools, 1983-1985;
Instructor for Measurement and Evaluation,
Oklahoma State University, 1984-1985; Research
Specialist, Oklahoma State Department of
Vocational and Technical Education, 1985-present.

Professional Organizations: American Educational
Research Association, American Vocational
Association, Kappa Delta Pi, Oklahoma
Vocational Association.