DOCUMENT RESUME

ED 195 981 CS 005 799

AUTHOR Byers, Joe L.; Evans, Thomas E.
TITLE Childrens' Reading Interests: A Study of Teacher Judgment. Research Series No. 81.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Sep 80
CONTRACT 400-76-0073
NOTE 34p.
AVAILABLE FROM The Institute for Research on Teaching, 252 Erickson Hall, Michigan State University, East Lansing, MI 48824 ($2.75)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Bibliographies; Childhood Interests; *Congruence (Psychology); Decision Making; Elementary Education; Identification (Psychology); *Reading Interests; *Reading Material Selection; *Reading Research; *Student Teacher Relationship; Teacher Attitudes; *Teacher Behavior

ABSTRACT

A total of 29 teachers and 227 elementary school students participated in a study that examined teachers' judgments of their students' reading preferences as indicated by book selections. Children were asked to choose books they would like to read from a list of book descriptions. Teachers were then asked to predict their selections. Results indicated that (1) the children's book selections were functionally related to the attributes of the book (cues) as identified from the book descriptions; (2) the teachers were using the cues in making their judgments; (3) although the teachers and their students were using the same set of cues to guide their judgments and selections, they often weighted the cues differently; and (4) there was evidence that cues not included in the cue list were functional for both students and their teachers. Although judgmental accuracy was, on the average, low, there was evidence that it could be improved. (HOD)

Research Series No. 81

CHILDRENS' READING INTERESTS:

A STUDY OF TEACHER JUDGMENT

Joe L. Byers and Thomas E. Evans

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

September 1980

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Education. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

INSTITUTE FOR RESEARCH ON TEACHING

Teachers' thoughts and decisions are the focus of studies currently under way at Michigan State University's Institute for Research on Teaching (IRT). The IRT was founded in April 1976 with a $3.6 million grant from the National Institute of Education. That grant has since been renewed, extending IRT's work through September 1981. Funding is also received from other agencies and foundations. The Institute has major projects investigating teacher decision-making, including studies of reading diagnosis and remediation, classroom management strategies, instruction in the areas of language arts, reading, and mathematics, teacher education, teacher planning, effects of external pressures on teachers' decisions, socio-cultural factors, and teachers' perceptions of student affect. Researchers from many different disciplines cooperate in IRT research. In addition, public school teachers work at IRT as half-time collaborators in research, helping to design and plan studies, collect data, and analyze results. The Institute publishes research reports, conference proceedings, occasional papers, and a free quarterly newsletter for practitioners. For more information or to be placed on the IRT mailing list please write to : The IRT Editor, 252 Erickson, MSU, East Lansing, Michigan 48824.

Co-Directors: Judith E. Lanier and Lee S. Shulman

Associate Directors: Lawrence W. Lezotte and Andrew C. Porter

Editorial Staff:

Lawrence W. Lezotte, coordinator of Communications/Dissemination
Janet Flegg, IRT editor

Abstract

This study examines the nature of teacher decision making in a realistic setting by looking at teachers' judgments of their students' reading preferences as indicated by book selections. The Brunswick Lens Model is used to describe and parse this judgment situation. Children were asked to choose books they would like to read from a list of book descriptions in a Scholastic Press brochure. Their teachers were asked to predict their selections. The results indicate that (1) the children's book selections were functionally related to the cues as identified from the book descriptions, (2) the teachers were using the cues in making their judgments, (3) although the teachers and their students were using the same set of cues to guide their judgments and selections, they often weighted the cues differently, and (4) there was evidence that cues not included in the cue list were functional for both students and their teachers. Though judgmental accuracy was, on the average, low, there was ample evidence that it could be improved.

CHILDRENS' READING INTERESTS: A STUDY OF TEACHER JUDGMENT

Joe L. Byers and Thomas E. Evans[1]

Two fourth-grade girls and one first-grade boy were asked to look at a set of brief descriptions of some books they might like to have. Their teachers were then given the same book descriptions and asked to judge (predict) which books each of their students had chosen. Correlations between the children's selections and their teacher's judgments were determined. The size of the correlations represent the judgmental accuracy of the teachers. The fourth-grade teacher had an accuracy of .1217 with one of her students and .5293 with the other. The first-grade teacher had an accuracy of .6966 with her student. These three cases show some of the variability that can be found in a judgment task like choosing books for elementary-school children.

Perhaps it is not surprising that there was a large discrepancy between the accuracy of judgment for the first fourth grader and the first grader; after all they were at different grade levels, of different sex, and had different teachers. Any or all of these factors may have contributed to the difference in accuracy. Studies of children's reading interests have repeatedly demonstrated that boys and girls don't share all the same interests, and that reading interests change as children mature. Perhaps the first-grade teacher simply knew more about the likes and dislikes of first-grade boys than the fourth-grade teacher knew about the likes and dislikes of fourth-grade girls. Because this

---

[1]Joe L. Byers is a senior researcher with the IRT, a professor of educational psychology, and director of the Office of Research Consultation, College of Education, Michigan State University. Thomas E. Evans is a former IRT research intern and an associate professor of psychology at Olivet College in Michigan.

explanation depends on general characteristics of the children, it cannot account for the discrepancy in the judgmental accuracy of the fourth-grade teacher between her two students. Factors within the judgment task itself may provide a more satisfying explanation.

A simple and therefore easily overlooked explanation of the differences in judgmental accuracy is that they are due to chance. If childrens' interests are inconsistent, then accuracy in prediction could only be a random factor. If this were the case, researchers would then want to know how consistently childrens' selections are determined by the nature of the books available for them to choose among.

Another possibility is that a teacher might not understand the nature of the judgment task. In any judgment task, the environment contains cues or attributes which are functionally related to the behavior of the people in the environment. In our setting the environment consists of the child and the books s/he is choosing among. The books differ along a variety of cues, such as type and theme. If teachers are unaware of what these cues are, or if they assume their pupils are using other cues that are irrelevant (e.g., reading level), then their judgments could not be expected to be systematically related to their students' choices. If the irrelevant cues being used by the teachers were correlated with relevant cues being used by the children, then accuracy may be greater than that due to chance. However, it could be negative if the relevant-to-irrelevant cue correlation were negative.

A variation on the theme of task understanding occurs when both teacher and child use the same cues but weight them differently. Such a circumstance could result in either low or negative judgmental accuracy.

The judgmental accuracy of children's interests varies widely. We suggest that the reasons for this may reside in the nature of the task itself, in the characteristics of the children and teachers involved, or both (Evans & Byers, Note 1).

It might be tempting to think that the entire problem of low accuracy could easily be avoided by simply allowing the children to select books for themselves. The purpose of the study was not to find an efficient way to choose books, but to examine the nature of teacher decision making in a realistic situation. Thus the discovery of large variations in accuracy of judgments provides researchers with conditions for learning about how teachers use the information available to them to made judgments and decisions.

The Lens Model

The Brunswick Lens Model has been successfully used to describe a wide variety of judgment tasks. It will be used here to parse the judgment situation we have set up. We wanted to see if it could help us understand what goes into a teacher's judgment of a child's reading interest, and if it could suggest ways these judgments might be improved.

A lens model has two components: a dynamic environment changing from state to state, and an observer or judge whose task it is to predict which of several states best describes the current environment. The current state of the environment is assumed to be functionally related to describable environmental characteristics, known as cues or attributes. The cues can be weighted and combined in a linear fashion to predict the current state of the environment. The judge faces the problem of discovering what the cues are in any given environment, which ones are relevant, and how weights are assigned and combined to produce the current state.

In our case, the model can best be understood by reference to Figure 1. The book selection task has three parts: a child, a set of book descriptions, and a teacher. The environment is the child's choice (or non-choice) of a book (the state) together with a set of cue values for description. The teacher does not know if the child has indicated a liking for a particular book (the state) but does have available the same cues as the child.

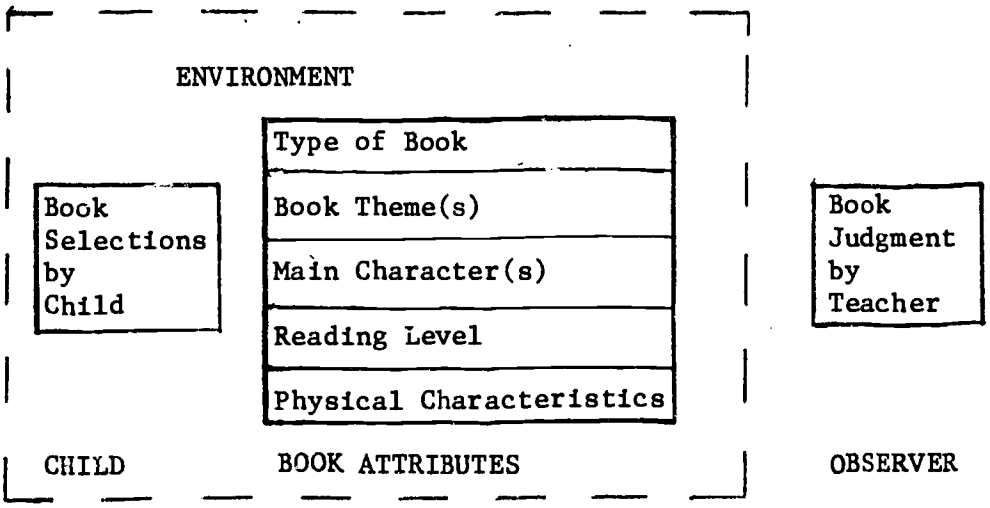

Figure 1. Schematic representation of book selection task as a lens model.

Figure 1 shows some of the broad classes of book characteristics which might be considered as cues in this judgment task. Scholars (Ford & Koplay, 1968; Oliver, 1977) have found that childrens' book preferences are consistently associated with variations in type of book (e.g., poetry, fiction, biography), themes (e.g., fantasy, heroism, survival), and main character (e.g., animal, parent, child, male, female). Additionally, childrens' preferences might be related to their reading level, or even to the physical characteristics of a book, such as its size, type font, or number of pages.

These same attributes (cues) of books could be expected to influence how teachers make judgments for a particular child's most likely selections. For example, if Janet is a child that likes animal stories, her teacher, upon seeing a list of books might muse, "Janet likes books about horses. I think she'd enjoy reading Black Beauty."

Thus the judgment task we are trying to understand has three components: (1) a set of attributes or cues of a book or set of books from which the choices are to be made, (2) the selections the child actually makes, and (3) the judgments or predicted selections made by his/her teacher. These parts are schematically represented in Figure 1.

The lens model for this judgment task consists of the three components just described. Given a quantification of these components, there are a number of useful scores which may be derived to shed light on the problem of the accuracy (or lack thereof) of the judgments.

First, as has already been indicated, given a child's selections and the teacher's predictions, the correlation between the two can be computed for any set of books. This correlation, the *judgmental accuracy* is the basic datum we are trying to understand.

Secondly, one can inquire as to how strongly the child's selections are determined by the book attributes we have listed. This inquiry follows multiple regression procedures in which the cue values for each of the books are used to predict the child's selection in much the same manner as high school rank, G.P.A., and SAT scores are used to predict high-school seniors' college performance.

The results of this analysis are critical to all that follows. The results are a multiple correlation coefficient that indicates the degree to which the child's book selections can be predicted from the characteristics of the books as quantified on the book cue list. This correlation is known as the *cue validity* or *ecological validity* of the

judgment task.

If the *cue validity* is low, then *judgmental accuracy* must also be low. In fact, usually *judgmental accuracy* cannot exceed the value of *cue validity*. If a child's book selections were random and not related to the values of the books on the book cue list, then no one could be expected to predict the selections. The *cue validity* was .7485 for the first fourth-grade girl; that's considerably higher than the accuracy of her teacher's prediction: .1217.

The teacher's judgment is also formulated in terms of the attributes (cues) of the books as well as *cue validity*. Regression analysis again allows a determination of the degree to which the judgments themselves are under the control of the values of the books on the book cue list. The multiple correlation coefficient derived from this analysis is known as the *cue utilization* or *cognitive control* of the judgment task. Generally speaking, the size of this variable sets the upper limit on *judgmental accuracy*. Suppose a child's book selections were perfectly predictable from the book cue list (*cue validity* = 1.00). If the teacher's judgment were capricious, say *cue utilization* equalled .2, then *judgmental accuracy* could not exceed .2.

Capriciousness is not the only thing that would produce a low *cue utilization*; it is not even a very feasible cause. A more plausible cause for low *cue utilization* is that the teacher is using attributes of books not explicitly considered in the model to help him/her make judgments. For instance, a teacher might not consider a given book because of its cost, or perhaps because of its controversial subject matter. The use of these attributes would result in a low *cue utilization* unless they were coded into the book cue list and thus made an explicit part of the judgment task. Children may also use cues not present in the model. For example, a child might make selections in terms of a book's

general popularity with his/her peers. This problem of using cues not explicitly contained in the model will be addressed shortly.

Suppose that for a given set of judgments there is both a reasonably high *cue validity* and *cue utilization*. What may then be predicted about *judgmental accuracy*? Under most circumstances, accuracy will not exceed the product of these two indices. For the first fourth-grade child, *cue validity* is .7485 and *cue utilization* is .6378. These values, though not high, are pretty good. Their product is .4774. Recall, however, that accuracy in this case was only .121, which suggests that another factor must be accounted for.

If accuracy is to be high, then not only must child and teacher systematically use the cues explicitly included in the model, but they must use them in the same manner. If a child picks books dealing with male characters and heroism while avoiding biographies, and his/her teacher perceives interests in science fiction and sports, there will be a mismatch in the weights assigned to the cues for the child and the teacher. Thus the equations for predicting the child's selections and the teacher's judgments would give rather different predictions. The correlation of the predictions from each regression (the child's and the teacher's) is the fourth index derivable from the lens model. This index represents the teacher's *knowledge* of the student's interest.

In a laboratory application of the lens model where both the child and the teacher were fully informed as to what the attributes (cues) were, the maximum value of accuracy would be strictly determined by the product of *cue validity, cue utilization,* and *knowledge*. In such a situation the difference between the product and the accuracy could be interpreted as a lack of fit of the model. Perhaps linear regression is not adequate to account for one or more of the indices.

In our application, the constraints necessary to allow a direct assessment of the formal aspects of linear regression would destroy the representativeness of this task. That is, it would no longer be a task or activity which one might find in an elementary school classroom. Since our main purpose here is not to extend lens-model theory but to use it in an attempt to become smarter about a phenomenon of classrooms, we must consider yet another index of the model.

If both child and teacher are using attributes of books that do not appear in the book cue list, and if these attributes were weighted similarly, then accuracy might be higher than the product of *cue validity, cue utilization,* and *knowledge* would demand. This last index we shall call *missing cue and/or non-linear cue weighting*; we'll call it *missing cues* for short. It is derived from correlating residuals from the child's selections with residuals from the teacher's judgments of those selections. If this correlation coefficient is neither zero nor close enough to zero that the discrepancy can be attributed to chance, it suggests that attributes are being used which are not included in the model.

We now apply our understanding of the five lens-model indices or parameters to the cases with which we began.

For the first fourth-grade student, there was a *knowledge* index of .367 and a *missing cue* index of -.128, along with *cue validity* and *cue utilization* of .653 and .765, respectively. Since *cue validity* is an index representing the predictability of the environment (i.e., the child and the book list), there is little the teacher could be expected to do about it. Though the *missing cue* value is not zero, it doesn't depart much from that value. There is little reason, therefore, to believe that the child and her teacher were using cues not on the book cue list when making their selections and judgments. This leaves two indices that are

potentially under the teacher's control: *cue utilization* and *knowledge*. Of these two, *knowledge* is the lowest. This suggests the teacher is not as knowledgeable about this student's reading interests as she might be, and that this fact has made a significant contribution to the low *judgmental accuracy* we earlier observed. If, for example, the teacher showed the same level of *knowledge* about this child as she did for the second fourth-grade student, the accuracy of the judgments would have reached approximately .31.

For the second fourth-grade student, the teacher achieved a *judgmental accuracy* of .5293. *Cue validity* and *cue utilization* were .7853 and .6997, respectively, while *knowledge* was .6527. The product of (.7853)(.6997)(.6527) is .3586, yet there is a higher accuracy than that. Thus the *missing cue* index is .3858. In this case, there is evidence that the teacher and the child share some cues that are not on the book cue list. The major results of the study will show how widespread this phenomenon is.

The teacher of the first-grade student achieved an accuracy of .6996, the highest of the three cases. *Cue validity* and *cue utilization* for this student were similar to those of the other two students. *Knowledge* was calculated at .7516, the highest value thus far encountered. Again, however, there is a rather high correlation between the residuals for the predictions of the child's choices and the teacher's judgments of those selections. This value (.6475) implies that both student and teacher were using cues not on the book cue list, or were combining cues in a complex and non-linear fashion.

The above has demonstrated how the lens model may be applied to a judgment task involving teachers making predictions about the books their students would like to read. These three illustrative cases show

(1) a wide range of *judgmental accuracy*, (2) that the children's choices were reliable (multiple R's in the .7 range),[2] (3) teachers' predictions were only slightly lower than that (R's = .6), (4) teacher knowledge of student interests varied widely, and (5) evidence that, for one teacher at least, cues not explicitly in the model accounted for a significant (no statistical inference intended) portion of *judgmental accuracy*. The remainder of this paper is addressed to reporting the findings of a large-scale study of teacher judgment of reading interests. The emphasis was on determining the average values of the lens-model parameters and assaying if they were influenced by the student's sex, grade level, and class membership (teacher effects).

Identification of Judgment Task Materials

A major priority of the study was to examine teacher's judgments of their pupils' reading interests in a natural setting. Given this, we decided to use optional reading materials readily available and widely used in mid-Michigan. Each month, Scholastic Press distributes a series of brochures describing a set of books which can be ordered by a teacher for his/her pupils. These brochures, printed by grade level (K-1, See Saw; grades 2-3, Lucky; and grades 4-6, Arrow) provide the book descriptions used in the judgment task. A rather brief, one or two sentence description follows the title of each book listed in these brochures. The childrens' selections and teachers' judgments were made from the brochures.

---

[2]Multiple-R represents the correlation between a child's book selections and the best prediction of those selections. It can range from 0.0 to 1.0. A 0.0 indicates no relationship between predicted and actual selections. A 1.0 means perfect predictability of actual selections from predicted selections. The value of 0.7 indicates that almost 50% of the variability in actual selections can be predicted.

Clearly, this source of attributes for describing books is not as complete as one derived from the books themselves or from a literary review of the books. Though we ran a risk of not capturing all the relevant attributes for the book cue list, we felt that to include more complete material would change the nature of the task because both teachers and pupils typically make their choices and judgments with only the brochure as a guide.

The development of the book cue lists followed several steps: literature review, analysis of childrens' verbal protocols, and finally the opinions of language arts experts.

Our extensive literature review of studies on reading preferences supported the following points: (1) Childrens' reading preferences are consistently related tc certain dimensions or features of books which can be reliably coded; (2) preferences are consistently associated with variations in type of character (animals, fantasy, peer group, parent, and child), type of activity (sex-related and age-related), and type of theme (realism, adventure/heroism, pranks and humor, and information); (3) patterns of preferences are likely to depend on the sex and age of the child; (4) although the dimensions are consistent, patterns of preference within these dimensions are highly individual and changeable; and (5) reading programs which include individualized interest patterns are likely to improve reading achievement and develop a more positive attitude toward reading for students (see Evans & Byers, Note 1, for a complete review of this literature).

The second step taken to identify important features of books was to analyze verbal protocols from student interviews. A group of 11 fourth-grade students (six boys, five girls) were presented with a

set of 36 books chosen from the Arrow series, 18 in each of two sessions. The children were asked to select from five to 10 books that they thought they would like and to describe to the interviewer why they liked each one. The interviews were recorded, transcribed, and analyzed for indications of the cue(s) used to make selections in either a positive or negative manner.

The third step was to elicit the opinion of expert judges. The literature review, protocols, and pupils' choices (plus the books used in the interviews) were presented to a panel of experts. This panel included researchers experienced in the study of judgment and experts in children's literature. After considerable deliberation, a list of cues was developed for pilot validation.

We attempted to validate our cues by selecting and coding a new sample of books along the dimensions determined by the expert panel. The new sample of book descriptions was presented to the same 11 fourth-grade children for them to choose from. The *cue validity* for the book cue list was computed for each of the children. The values ranged from a low of .516 to a high of .727. Further, the book cue intercorrelations were all quite small (little co-linearity).

The final list of book cues is presented in Table 1. This table also shows the range of books that were included in the major study, and may, therefore, give the reader some notion about how the publisher (Scholastic Press) believes childrens' reading preferences change as they progress through elementary school. For the final book list each book was independently rated by at least four or five professionals. Their ratings were summed and averaged, which gave the book cue list an interval-scale-like property. The value of a cue for any particular book could range from zero (no raters believed it present) to 1.0, indicating all raters agreed that cue was present.

Table 1.

Cue Frequency Expressed as a Percentage of

Books for Three Lists Covering

Grades 'K' to Sixth

Types of Cues	Kindergarten & First Grade (N=38)	Second & Third Grade (N=50)	Fourth & Sixth Grade (N=91)
1. Type: Biography	0.00	12.00	10.48
2. Info: Science	23.07	22.00	8.87
3. Info: History	0.00	4.00	4.03
4. Info: Other	0.00	10.00	24.19
5. Fantasy	69.23	42.00	20.96
6. Science Fiction	0.00	2.00	7.25
7. Folktales, Fables, Myths	25.64	12.00	4.03
8. Historical Fiction	0.00	10.00	9.67
9. Realistic Fiction	41.02	38.00	50.00
10. Poetry	0.00	2.00	0.80
11. Picture Book	56.41	34.00	11.29
12. Themes: Coping	43.58	20.00	47.58
13. Danger	46.15	32.00	45.96
14. Devotion	23.07	16.00	25.00
15. Heroism	25.64	20.00	27.41
16. Home and Family	35.89	42.00	24.19
17. Humor	64.10	50.00	29.03
18. Sports	0.00	6.00	12.90
19. Horrible and Hideous	35.89	10.00	10.48
20. Death	0.00	4.00	9.67
21. Drugs and Alcohol	0.00	0.00	0.80
22. Mystery and Surprise	51.28	46.00	41.93
23. Survival	33.33	30.00	35.48
24. Male Character	64.10	30.00	50.80
25. Female Character	35.89	38.00	37.09
26. Animal (not horses)	74.35	52.00	27.41
27. Horses	0.00	8.00	12.90
28. Significant Others	33.33	36.00	28.22
29. Peers	0.00	18.00	20.96

Method

Subjects

A total of 29 teachers and 227 students participated in the study. The distribution of teachers and students in terms of grade level and sex of students and location of classrooms is given in Table 2. Students were selected at random from class lists prepared by the teachers, subject to the restriction that there should be an equal number of boys and girls.

Materials

The materials used in this study have already been described in this report. It only remains to note that for a given set of books, a cue was dropped from the book cue list if it was not rated present by at least one rater in at least five of the books on that list. This rule resulted in reducing the size of the book cue lists from a potential of n (number of books) rows and 29 columns (one for each cue) to smaller sizes. Table 2 shows the actual size of each of the book cue lists used for the various sample groups in the study.

Procedure

Students were taken in small groups (6-10) to a quiet room in their school. The researcher distributed the brochures to students and told them that he was interested in what books they liked to read. With the kindergarteners, the researcher read the book descriptions to the children and then helped them place little plastic stickers over the titles of the books they thought they'd like to have. Older children were given sheets with the titles printed on them and asked to check those they'd like to have. In all cases, children were encouraged to select not more than 10 books from their lists. Data collection occurred on two

Table 2.

Numbers, Grade Level, Location, and Sex of Children and Teachers Used in the Study of Teacher Judgments of Children's Reading Preferences

Grade Level	Teachers		Number of Children			Number of Cues Used	Book List Length	Book Selection Decisions
	No.	Location	Boys	Girls	Total			
Kindergarten	2	Lansing	4	4	8	18	39	312
First	7	Lansing	20	19	39	18	39	1521
Second	10	Lansing	43	40	83	19	50	4150
Fourth	4	Lansing	19	20	39	25	91	3549
	1	Olivet	5	5	10	25	75	750
Fifth	2	Olivet	10	10	20	25	75	1500
Sixth	3	Jackson	14	14	28	17	42	1176
Totals	29		115	112	227			12958

occasions about four-to-five weeks apart. In order to get a large enough sample size for each student/teacher pair, the data from the separate occasions were analyzed together. After we had collected data from the students, we asked their teachers to judge the selections each of their students had made. They had no information about what selections had actually been made.

Results

The final data set for analysis consisted of 227 cases, each with five scores representing the lens-model parameters. This final data set had to be prepared by analyzing each student/teacher pairing by lens-model procedures. Programs were written in APL (A Programming Language) to take the following as input: a student's book selections, the book cue list used, and the teacher's judgments of the students' choices. These programs returned scores for the five parameters (*judgmental accuracy, cue validity, cue utilization, knowledge,* and *missing cues*) as output.

The analysis of the final data consisted of three stages. The first stage focused on estimates of the five parameters. Stage two focused on components of variance analysis that examined three factors likely to contribute to variability in the values of the parameters. The final analysis was a multivariate analysis of variance that examined the specific influence of the student's grade level and sex on the parameters of the lens model.

Stage 1

We felt that because of the large number of students and teachers from various elementary-school grade levels and locations, the absolute size of the estimated values for the parameters should give a reasonably stable population value for this type of judgment task. Table 3 gives the means, standard deviation, and 95% confidence intervals for each

parameter.

Table 3.

Estimates of Lens-Model Parameters

Parameter	Mean	Standard Deviation	95% Confidence Interval Lower Limit	Sample Value	Upper Value
Judgmental Accuracy	227	.1861	.20763	.23431	.26099
Cue Validity	227	.0941	.63755	.67324	.70893
Cue Utilization	227	.0928	.64061	.67862	.71663
Knowledge	227	.2434	.23333	.31161	.38988
Nonlinear/missing cues	277	.2255	.11889	.16077	.20264

Judgmental accuracy for this task varied between .21 and .26. The point estimate was .23, or that 5.5% of the variability in the pupil's selections could be predicted from the teacher's judgments. The accuracy of judgment is a function of the remaining parameters, and, therefore, any final interpretation of this rather poor showing should be postponed until the other values have been discussed.

Cue validity represents the degree to which pupils are making book selections in a predictable manner. The 95% confidence interval for *cue validity* is .64 to .71. This provides evidence that selecting books from the book descriptions found in the Scholastic Press brochures was able to capture a significant portion of the variance (45%) in pupils' reading preferences.

Cue utilization, or the degree to which teachers used the cues in making their judgments of their pupils' reading interests was about the same magnitude as *cue validity* (see Table 3). Thus teachers found the task to be one that they could respond to in a systematic fashion. *Cue utilization*, unlike *cue validity*, can be improved. The judgment literature

reports procedures for increasing *cue utilization* (sometimes referred to as cognitive control). For instance, Hammond and Summers (1972) and their colleagues demonstrated that process feedback[3] could significantly improve *cue utilization*. Naturally, a procedure which results in the improvement of *cue utilization* will also elevate *judgmental accuracy*.

Table 3 reveals that *knowledge* is a very important factor in the observed low level of *judgmental accuracy*. Improvements in *knowledge* would result in improvements in *judgmental accuracy*, especially if coupled with an increase in *cue utilization*. Finally, there was evidence that pupils or teachers or both used cues not included in the book cue list (Table 1), or devised non-linear weighting schemes to direct their selections and/or predictions.

Stage 2

The results of analysis in Stage 2 focus on factors which may contribute significantly to the variance of the five parameters. Since a purpose of the study was to collect data in a realistic setting, it was not possible to balance the design and cross all the factors of interest. Three factors were examined for their potential influence on the lens-model parameters; grade-level, the individual differences between teachers, and sex of pupil. These factors, for Stage 2 analysis, were treated in a nested unbalanced design in which pupils were nested within sex of pupil (Factor 1), sex of pupil was nested within individuality of teachers (Factor 2), and individuality of teachers was nested within grade-level (Factor 3). The analysis followed procedures established by

[3]Hammond & Summers (1972) make a distinction between outcome feedback, how accurate a judgment is, and process feedback, how the judge uses (thus *cue utilization*) or weights the cues in arriving at a judgment. They demonstrated that process feedback was more effective than outcome feedback in influencing *judgmental accuracy*.

Garrett & Goss (1980), which yielded estimated variance components, the percent of variance for each component, and an F-test for the significance of the component. Table 4 presents the results of this analysis.

For *judgmental accuracy*, grade-level contributed no variance to the total, individual differences between teachers accounted for 34.58%, sex of pupil for 6.85%, and pupils for 58.57%. F-tests revealed that the variance component attributed to teachers' individual differences was significantly different from zero ($F = 4.15$; $p < .01$). Different teachers had a significant influence on *judgmental accuracy*. Other than the teacher component, no other source of variance was found to be reliable.

Grade-level was shown to be a significant source of variance for both *cue validity* and *cue utilization*. Sex of pupil was a contributing factor to *cue utilization*. Interestingly, differences between teachers did not contribute to the variance of either of these two measures.

Teachers' individual differences were a significant source of variance for *knowledge* and for *missing cues*, while grade-level and sex of pupil were not. Of the three factors considered in the components of variance analysis, only the individuality of teachers could be considered as a random factor. The influence of this random factor is found primarily in *judgmental accuracy*, *knowledge*, and *missing cues*. The fact that *cue validity* was not influenced by teachers' individual differences and was found to be related to grade level, helps to confirm the belief that the lens-model approach to the study of teacher judgment is a reasonable one. *cue validity* is not directly related to the individuality of teachers, but could be argued to be related to the age and/or development of the pupil. Thus the relationship of *cue validity* to grade-level is understandable. This relationship will be discussed further under Stage 3.

Table 4.

Components of Variance for Five Lens-Model Scores Using a Nested Unbalanced Design: Pupils within Sex of Pupil, Sex of Pupil within Teacher's Individual Differences, and Teachers' Individual Differences within Grade Level

Lens-Model Scores	Source of Variance				
Components of Variance (Degrees of Freedom)	Grade Level (5)	Teachers (23)	Sex of Pupil(s) (29)	Pupils (169)	Total (226)
Judgmental Accurracy					
Mean Square	.056323	.13036	.031314	.021584	
Estimated Variance	-.00217	.012713	.0025185	.021534	.036765
Percent of Total	0.0	34.58	6.85	58.57	100.00
F-Value		4.15**	1.45		
Cue Validity					
Mean Square	.032113	.010386	.0099882	.0077749	
Estimated Variance	.00062388	.000048117	.0005699	.0077749	.0090168
Percent of Total	6.92	0.53	6.32	86.23	100.00
F-Value	3.08*	1.04	1.28		
Cue Utilization					
Mean Square	.035084	.0077097	.015179	.068209	
Estimated Variance	.00078782	-.00098114	.0021523	.0068209	.009761
Percent of Total	8.07	0.0	22.05	69.88	100.00
F-Value	4.56**		2.23***		
Knowledge					
Mean Square	.19050	.16845	.050956	.041948	
Estimated Variance	.00057794	.015085	.0023195	.041948	.059931
Percent of Variance	0.96	25.17	3.87	69.99	100.00
F-Value	1.12	3.30**	1.21		
Non-Linear and Missing Cues					
Mean Square	.078351	.12318	.54352	.039620	
Estimated Variance	-.13252	.0088245	.0037936	.039620	.052239
Percent of Variance	0.0	16.89	7.26	75.84	100.00
F-Value		2.26*	1.37		

* $p < .05$
** $p < .01$
*** $p < .001$

Stage 3

The final analysis explored the effects of grade-level and sex of pupil on the parameters of the lens model. Grade-level and sex of pupil were fixed factors and were fully crossed yielding a three by two factorial multivariate analysis of variance. Table 5 presents the mean and standard deviation for each dependent variable as well as the cell sizes. The design, though crossed, was unbalanced due to varying cell sizes. This meant that the tests for main effects and interactions were not independent and required the inspection of the grade-level by sex of pupil interaction first.

Since this analysis uses only grade-level and sex of pupil and ignored teachers' individual differences, the variance due to this last factor remained in the residual estimates of variance and covariance. The effects of this are revealed in the within-cells correlations of the lens-model parameters (Table 6). Both *missing cues* and *knowledge* were found to correlate about .79 with *judgmental accuracy*. Recall that Table 4 showed teachers' individual differences to be a significant source of variance for *judgmental accuracy, knowledge,* and *missing cues*.

The grade level by sex of pupil interaction yielded a significant multivariate test (Lambda = .7745; F = 2.24; d.f. = 25/785.3; $p < .001$). Thus formal hypothesis testing stopped at this point. Main effects of grade level or sex of pupil would be confounded by the existence of the grade level by sex of pupil interaction.

The grade level by sex of pupil interaction was scrutinized further by means of step-down F-tests in an attempt to pinpoint the locus of the interaction effect. The parameters were ordered as follows for the step-down tests: (1) *missing cue,* (2) *cue validity,* (3) *cue utilization,*

(4) *knowledge*, and (5) *judgmental accuracy*. This ordering implied a causal model in which *judgmental accuracy* was examined for interaction effects after all the variance due to the other parameters had been removed. The data revealed that there was no interaction effect for *judgmental accuracy*. ($F = 1.84$; d.f. = 5/211; $p < .11$). The next parameter to be examined was *knowledge*. For this dependent variable, the covariance to *cue validity*, *cue utilization*, and *missing cues* was removed prior to testing for evidence of interaction effects. Again the data did not lead to a rejection of the null hypothesis ($F = 0.76$; d.f. = 5/212; $p < .57$). Grade level and sex of pupil did not interact to produce any reliable difference in the sample means for *knowledge*.

Missing cues and *cue validity* covariance was removed from the variance of *cue utilization*. The conditional *cue utilization* variance was then tested for the grade level by sex of pupil interaction, and this time the test was reliable ($F = 4.93$; d.f. = 5/213; $p < .001$). Figure 2 presents the grade level by sex of pupil interaction graphically for *cue utilization*. It is important to recognize that the scaling of this graph may exaggerate the differences between boys and girls or the changes over grade level. With that caveat in mind we will speculate a bit. Taken as a whole, the pattern of *cue utilization* for girls is cup shaped. It starts at the kindergarten level (0.0) near .76 and decreases through the fourth grade. Thereafter it rises. On the other hand, the boys' *cue utilization* profile was one of general decline over the period of grades included in the study.

Cue validity was conditioned only on *missing cues*. Here, as with *cue utilization* the interaction test was significant ($F = 2.81$;

Table 5.

Average Values for the Lens Model

For Boys And Girls In Grades K-6

		Boys			Girls		
		Mean	S.D.	N	Mean	S.D.	N
Kindergarten	JA	0.292	0.055	4	0.223	0.367	4
	CV	0.619	0.122	4	0.769	0.083	4
	CU	0.749	0.129	4	0.755	0.120	4
	K	0.287	0.192	4	0.181	0.469	4
	C	0.248	0.121	4	0.231	0.212	4
First Grade	JA	0.193	0.218	20	0.279	0.192	19
	CV	0.700	0.088	20	0.730	0.087	19
	CU	0.663	0.074	20	0.779	0.071	19
	K	0.257	0.260	20	0.320	0.289	19
	C	0.137	0.274	20	0.193	0.259	19
Second Grade	JA	0.254	0.178	43	0.260	0.194	40
	CV	0.680	0.068	43	0.673	0.070	40
	CU	0.681	0.065	43	0.672	0.063	40
	K	0.327	0.221	43	0.327	0.252	40
	C	0.186	0.211	43	0.198	0.242	40
Fourth Grade	JA	0.234	0.151	24	0.263	0.163	25
	CV	0.620	0.080	24	0.647	0.156	25
	CU	0.647	0.065	24	0.638	0.144	25
	K	0.386	0.221	24	0.429	0.193	25
	C	0.126	0.145	24	0.131	0.172	25
Fifth Grade	JA	0.190	0.182	10	0.077	0.123	10
	CV	0.650	0.055	10	0.652	0.064	10
	CU	0.675	0.094	10	0.654	0.104	10
	K	0.229	0.231	10	0.212	0.202	10
	C	0.141	0.216	10	0.024	0.168	10
Sixth Grade	JA	0.186	0.211	14	0.227	0.182	14
	CV	0.626	0.072	14	0.742	0.082	14
	CU	0.615	0.062	14	0.745	0.076	14
	K	0.115	0.250	14	0.323	0.155	14
	C	0.225	0.257	14	0.118	0.299	14

Key

JA = Judgmental Accuracy
CV = Cue Validity
CU = Cue Utilization
K = Knowledge
C = Missing Cues

Table 6.

Within Cells Correlations (S.D. on Diagonal)
of Lens-Model Parameters

	Missing Cue	Cue Validity	Cue Utilization	Knowledge	Judgmental Accuracy
Missing Cue	(.2247)				
Cue Validity	-.0527	(.0885)			
Cue Utilization	-.0567	.3019	(.0840)		
Knowledge	.2843	.2091	.1832	(.2368)	
Judgmental Accuracy	.7908	.1215	.11056	.7920	(.1853)

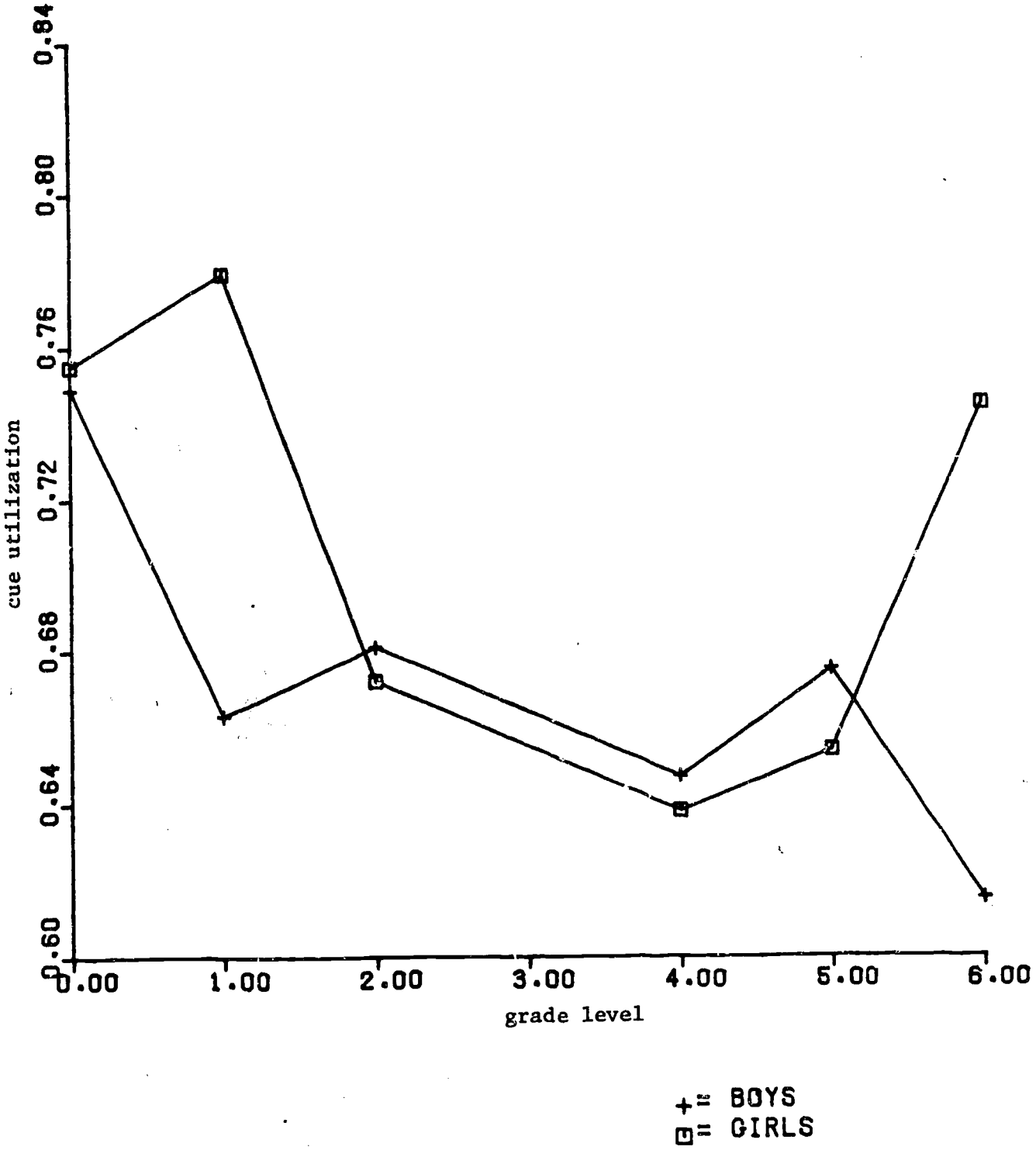

Figure 2. Average values of *cue utilization* for boys and girls across grades K-6.

d.f. = 5/214; $p < .02$). Figure 3 presents the grade by sex interaction for *cue validity*. Again, as for *cue utilization* care must be taken to avoid the impression of very large differences. The pattern for *cue vallidity* was quite similar to that found for *cue utilization*.

Discussion and Conclusions

The main analyses of the study focussed on estimating population values for the five parameters of the lens model (*judgmental accuracy, cue validity, cue utilization, knowledge,* and *missing cues.)* and identifying factors that might contribute to variability in their value. The results of the estimation stage led to the conclusion that though *judgmental accuracy* was, on the average, low, there was ample evidence that it could be improved. First, *cue validity* was about .6, indicating that the pupil's selections were functionally related to the cues as identified from the book descriptions. Second, *cue utilization* averaged about .6, which indicated that teachers were using the cues in making their judgments. Thus appropriate training could be expected to increase *cue utilization*. Thirdly, *knowledge* was low, about .3; thus although both teachers and their pupils were using the same set of cues to guide their judgments and selections, they often weighted the cues differently. Lastly, there was evidence that cues not included in the cue list were functional for both pupils and their teachers (See Table 3).

A components-of-variance analysis and a multivariate analysis of variance revealed that *knowledge* and *missing cues* were most strongly related to *judgmental accuracy*. They further revealed that neither grade-level nor sex of pupil were significant sources of variance for these variables. Teachers' individual differences contributed to

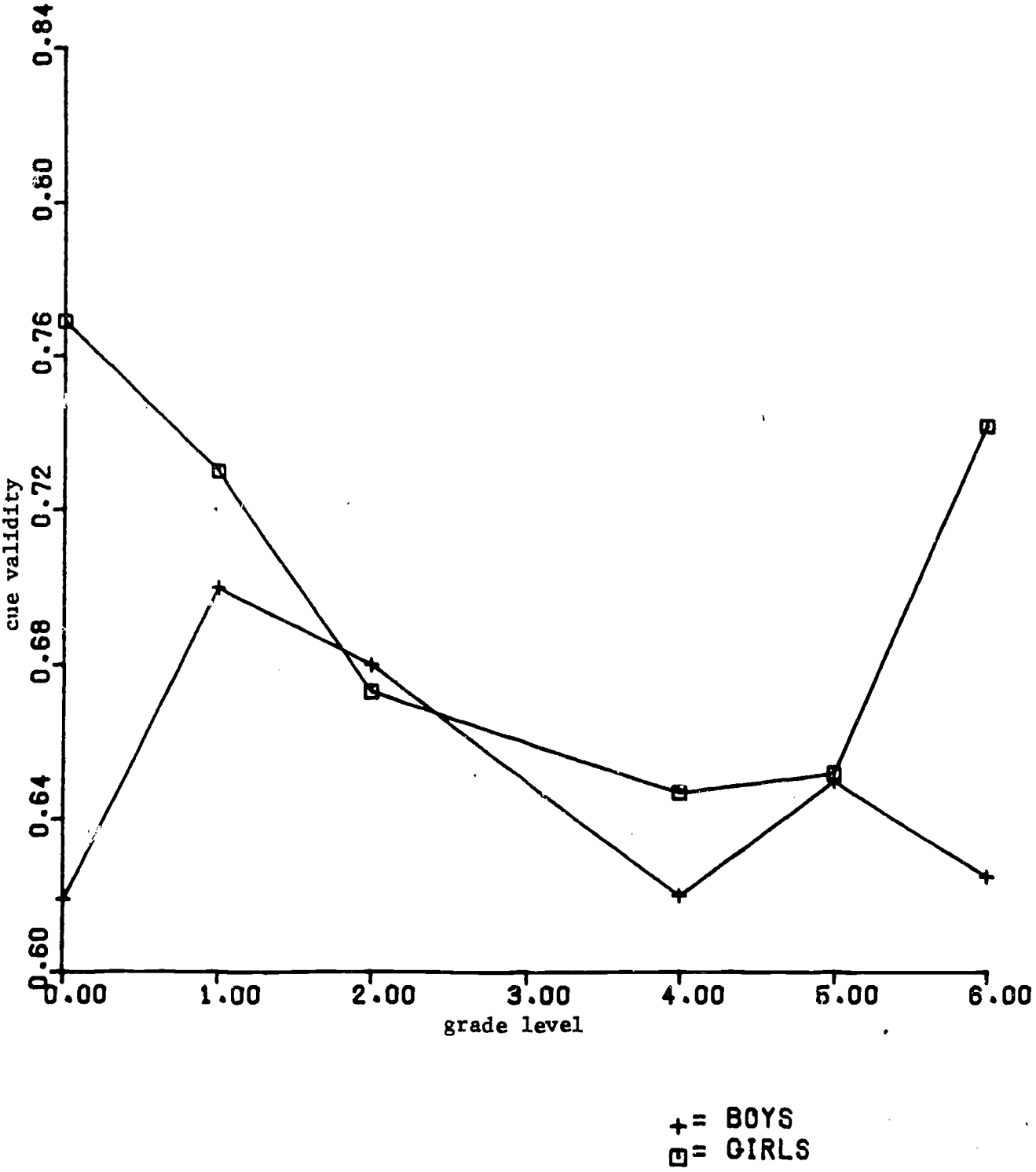

Figure 3. Average values of *cue validity* for boys and girls across grades K-6.

differences in *judgmental accuracy, knowledge,* and *missing cues*. These results may be interpreted to mean that the development of sex-bias phenomena did not influence the accuracy of judgments in this task.

The meaning of the grade level by sex of pupil interaction involving *cue validity* and *cue utilization* are somewhat more difficult to interpret. It may be that girls' reading interests at lower elementary grades are fairly stable and well predicted by the cue list. As their reading skill increases, their interests may become less predictable until the middle grades, when they begin to stabilize again. It is possible that the pattern for boys is similar to that for girls, but is simply delayed for a year or two. Thus the current data for boys showed only the first part of the profile (i.e., the declining part).

As noted above, *cue utilization* followed a similar pattern, over time, to *cue validity*. No readily apparent explanation for this comes to mind unless teachers are trying to track the developmental changes of boys and girls' interests.

Reference Notes

1. Evans, T. E., & Byers, J. L. Teacher judgment of children's reading preferences (Research Series No. 38). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

References

Ford, R. C. & Koplyay, J. Children's story preferences. Reading Teacher, December 1968, 22, 233-237.

Hammond, K. R., & Summers, D. A. Cognitive control. Psychological Review, 1972, 79, 58-67.

Oliver, L. The reading interests of children in the primary grades. The Elementary School Journal, 1977, 77, 401-406.