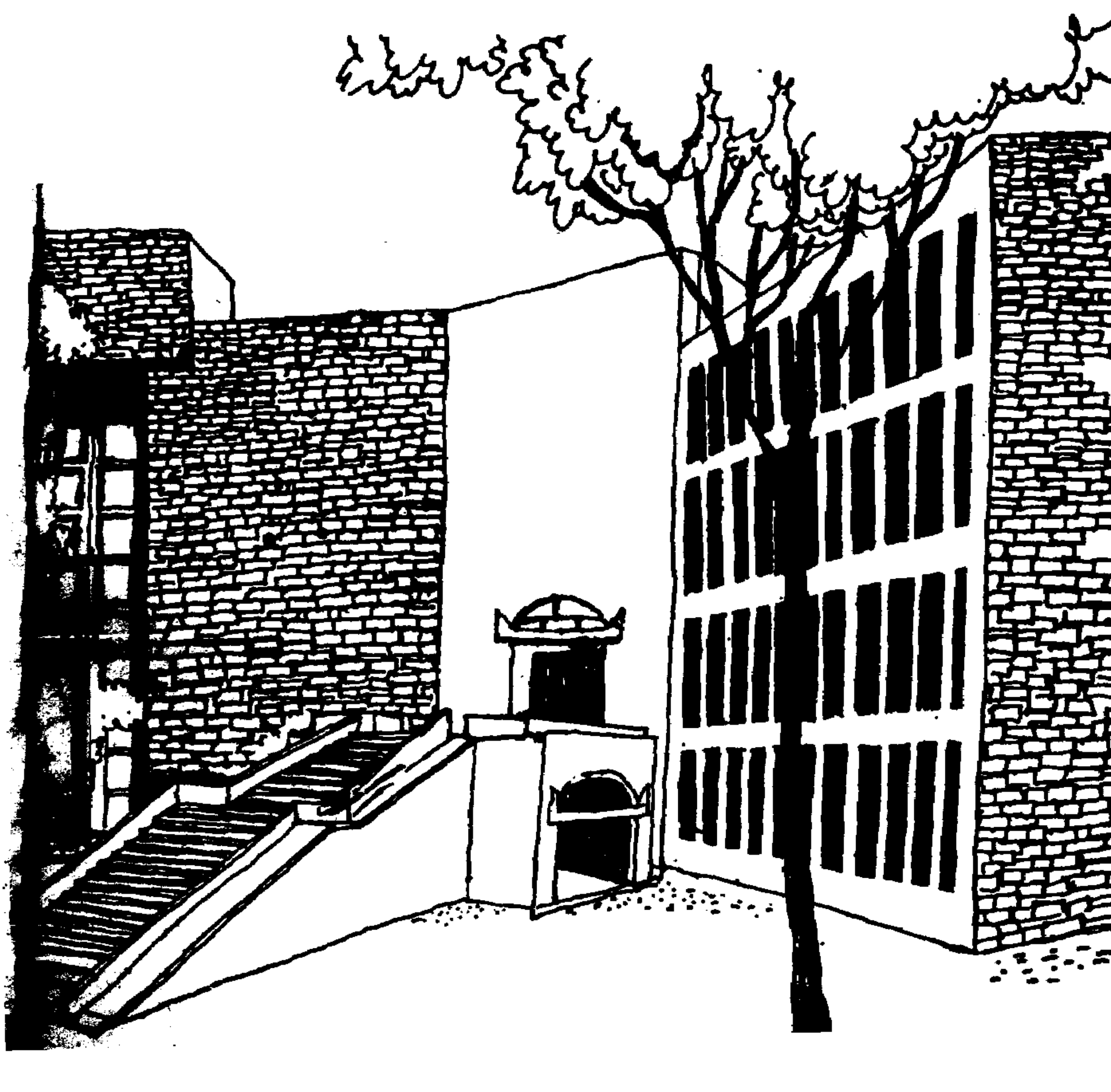
IIM AHMEDABAD

W.P. 634

Working Paper



DETERMINANTS OF ERROR VARIANCE IN INDIVIDUAL CHILD ANALYSES FOR STUDYING INTEGRATIONAL CAPACITY

By

Ramadhar Singh

W P No. 634

October 1986

The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380015
INDIA

Determinants of Error Variance in Individual Child Analyses for Studying Integrational Capacity

Ramadhar Singh

Organizational Behavior Area

Indian Institute of Management, Ahmedabad, India

Running Head: ERROR VARIANCE

Abstract

In research on information integration by children, integrational capacity of a child is ascertained by the number of statistically significant main effects in analysis of variance for the child. Such an approach has an obvious bias, for the error variance may be bigger for younger but smaller for older children. Analyses of error variance from individual child analyses of four previous experiments on prediction of performance yielded mixed results. The mean error term sometimes decreased, sometimes increased, and sometimes remained the same over increasing age of children. The number of cues in the integration task and the nature of judgment seemed to be better predictors of error variance than the age of children. It seems that a study of integrational capacity must as a rule check on response variability along with the number of statistically significant main effects over ages.

In research on information integration by children, integrational capacity of a child is estimated from the number of pieces of given information that the child utilizes (Anderson, 1980). Each child judges the same set of stimuli constructed from a factorial design at least over two trials of judgment. Integrational capacity of a child is ascertained by the number of statistically significant main effects in analysis of variance for the child. The greater the number of statistically significant main effects, the greater is the integrational capacity of the child (Anderson & Butzin, 1978; Gupta & Singh, 1981; Kun, Parsons, & Ruble, 1974; Leon, 1982; Miller, 1982; Srivastava & Singh, 1986a, 1986b).

As the statistical tests of main effects in the individual child analyses use the pooled within-cell variability over trials of judgment as the error term (Anderson, 1982, p. 64), there is an obvious bias in accepting statistically significant main effects as direct indices of integrational capacity. On the basis of common sense, the response variability may be expected to decrease over increasing age of children. If this is correct, then the error term for the tests of significance would be larger for younger children but smaller for older children. As a result, the main effect of the same magnitude could be statistically significant for an older but not for a younger child (Anderson & Butzin, 1978).

The purpose of the present report is two-fold. One purpose is to present new data on the relationship between error variance in individual child analyses and age of children. Anderson and Butzin (1978, p. 602) found a steady decrease in response variability of 4- through 8-year-olds, but reliability of this age-trend has not been checked in data from other developmental studies. Another purpose is to examine the effects of the number of cues and of nature of judgmental task on error variance. If response variability is a genuine age-related problem perhaps attributable to variable integrational capacity, then error variance should be smaller for a simple, two-cue task than for a complex, four-cue task. No such test has so far been reported.

'o accomplish the two purposes just mentioned, mean error variance from the individual subject analyses of the four experiments of the three previous studies of prediction of performance from information about motivation and ability of stimulus students (Gupta & Singh, 1981; Srivastava & Singh, 1986a, 1986b) were examined. The effects of age of subjects and of the number of cues and nature of judgmental task on error variance were of main interests. A brief description of the method of each of the four experiments and the results pertaining to error variance are presented below.

Gupta and Singh, 1981

Method

Ninety-six children and 24 adults predicted exam performance of hypothetical students on nine squares, arranged in their increasing size from 1 to 9 cm. Some stimulus students were described with respect to their performance in the previous grade as well as their current motivation and ability; others were described with respects to their current motivation and ability. The three-cue descriptions were prepared from a 3 x 3 x 3 (Past performance x Motivation x Ability) factorial design; the two-cue ones were from a 3 x 3 (Motivation x Ability) factorial design.

Each child judged all the three-cue and two-cue descriptions presented in thoroughly shuffled orders over two days. Although 18 practice examples were used before data collection, judgments from the first presentation of the experimental stimuli were considered as additional practice for the subject. Only the data from the two trials of judgment on the second day were analyzed. Each adult rated the stimuli three times on the same day. Data from the last two trials of judgment were analyzed.

The children were from Standard II, IV, VI, and VIII of the Campus School and the Central School of the Indian Institute of Technology, Kanpur, India. There were 12 boys and 12 girls in each grade group. Mean ages for the four grade groups were 6 years 9 months, 8 years 7 months, 10 years 7 months, and 12 years 9 months, with respective ranges of 6-7 years, 8-9 years, 10-11 years, and 12-13 years. Each child received five toffees for his or her cooperation in the experiment.

The 24 adults (12 men and 12 women) were undergraduate engineering students at the Indian Institute of Technology, Kanpur. The mean age was 17 years 10 months, with a range of 17-19 years. Participation fulfilled the s students' course requirement.

<u>Results</u>

The mean error term for the subjects of 6-7, 8-9, 10-11, 12-13, and 17-19 years were 2.17, 0.87, 0.77, 0.88, and 0.55, respectively. The overall effect of age of subjects was highly significant, $F(4, 115) = 18.64$, $p < .01$. Post-hoc comparisons among means by the Newman-Kauls test (Winer, 1971) at .05

level of significance further disclosed that 6-7-year-olds had bigger error term than the other four groups of subjects who did not differ among themselves.

This result of a greater response variability in 6-7-year-olds provides an independent replication of the finding of Anderson and Butzin (1978). No less important, the present analyses show that response variability is not a problem in developmental research with children of 8-9 years and above.

The mean error term for the judgments of the three-cue (*M* = 1.04) and two-cue (*M* = 1.08) descriptions did not differ at all, *F* (1, 115) = 0.00. The interaction between age of subjects and number of cues in descriptions was not significant either, *F* (4, 115) = 0.74. This null result implies either that the two-cue and three-cue descriptions perhaps did not differ substantially in their complexity or that response variability is not linked with the complexity of task.

Srivastava and Singh, 1986a

Method

One hundred and 20 kindergarten through fourth standard children from the same population as in Gupta and Singh (1981) predicted exam performance of hypothetical students along a 21-step ladder. Student descriptions contained information about current motivation and ability, and each kind of information supposedly came from two independent sources. These four-cue descriptions thus constituted a four-opinion integration task for the children.

There were two-cue descriptions also. They contained either motivation or ability information from the same two different sources as in the four-cue descriptions.

The four-cue descriptions were prepared from a 2 x 2 x 2 x 2 (Mother x Neighbor x $Teacher_1$ x $Teacher_2$) design. The first two sources supplied information about the time spent by the stimulus student over study; the last two sources gave opinions on student's potential for learning. The two two-cue designs were 2 x 2 (Mother x Neighbor) and ($Teacher_1$ x $Teacher_2$) factorials.

Each child judged all the four-cue and two-cue descriptions presented in thoroughly shuffled orders over three consecutive days. After working with 14 practice examples, the child rated the stimuli one time on the first day and two times on both the second and third days. Judgments from the first day were treated as additional practice just as in Gupta and Singh (1981). Only the data from the two trials of the second and third days were analyzed.

Subjects were kindergarten, first, second, third, and fourth standard children. There were 12 boys and 12 girls in each age group. Mean ages for the five groups of children were 4 years 6 months, 5 years 10 months, 7 years, 8 years, and 9 years 2 months with the respective ranges of 4 years 3 months to 5 years 3 months, 5 years 5 months to 6 years 4 months, 6 years 5 months to 7 years 6 months, 7 years 7 months to 8 years 9 months, and 8 years 6 months to 10 years 2 months. Each child received five toffees and five balloons for his or her cooperation in the experiment on each day.

This experiment differed from that of Gupta and Singh (1981) in at least four major ways. First, although there were four pieces of information available for judgment, they were of only two kinds—motivation and ability. As already noted, Gupta and Singh had used three qualitatively different kinds of information, namely, past performance, motivation, and ability. Second, the number of trials of judgment were four, twice more than those in Gupta and Singh. Third, the response measure was a 21-step ladder, not a set of nine squares which varied in both the length and width. Finally, children were motivated to take the task seriously by giving five toffees and five balloons on each day.

Results

Table 1 presents mean error term as a function of the number of cues in the judgmental task and age of children. Two results stand out most clearly from Table 1. First, mean error term is the highest for children of Standard IV which qualifies the conclusion drawn from the previous analyses. Second, the mean error term is uniformly less for the four-cue task than for the two-cue tasks. This result is contrary to the hypothesis that lack of integrational capacity would cause higher error variance in complex than in simple task.

Table 1 about here

Both results were supported by statistical analyses also. The mean error term averaged over the three tasks were 1.13, 0.86, 0.97, 0.85, and 1.54 for kindergarten, first, second, third, and fourth standard children, respectively. There was a statistically significant effect of age of subjects, F (4, 115) = 3.17, $p < .05$, and children of fourth standard had significantly higher mean error term than those of first, second, and third standards according to the Newman-Keuls tests. Furthermore, the mean error term for kindergarten children did not differ from any of the other four mean error terms.

The mean error term averaged over the subjects of five age groups were 0.82, 1.11, and 1.28 for the judgments of descriptions constructed from Mother x Neighbor x $Teacher_1$ x $Teacher_2$, Mother x Neighbor, and $Teacher_1$ x $Teacher_2$ designs. There was a statistically significant main effect of the number of cues in task, F (2, 230) = 15.94, $p < .01$, and post-hoc comparisons disclosed that the mean error term of the four-cue task was significantly less than those of the two two-cue tasks. The mean error terms of the two two-cue tasks also differed significantly.

The foregoing results suggest that error term in the individual child analyses cannot be linked directly with the age of children. Age produces an effect on response variability but there is no evidence for a steady decline in response variability over increasing age as Anderson and Butzin (1978) found. It is also clear that simple tasks engender greater response variability than do complex ones, and that what enter as inputs in the task are major determinants of error variance in individual child analyses.

Srivastava and Singh, 1986b

<u>Experiment 1</u>

<u>Method</u>. Students of Standard IV and VIII predicted performance of 10-year-olds who were to participate in either a puzzle or singing contest to be organized in the school. Profiles of nine stimulus students were prepared from a 3 x 3 (Motivation x Ability) factorial design. Each kind of information was

shown by a series of seven vertical bars, having width of 1 cm but height from 1 to 7 cm. The levels of the two factors were second, fourth, and sixth vertical bars.

Each subject looked at the motivation and ability of the stimulus student and predicted his or her contest performance along the 21-step ladder. The same set of stimuli were presented three times in different shuffled orders. Data from only the second and third presentations of the stimuli were analyzed.

Ninety-six children from the Central School, Shahibaugh, Ahmedabad, India, served as subjects. They were selected according to a 2 x 2 x 2 design, having standard of subjects (Standard IV versus Standard VIII), sex of subjects (boy versus girl), and nature of contest (puzzle versus singing) as factors. There were 12 subjects in each of the eight cells. The mean ages for the children of Standard IV (n = 48) and VII (n = 48) were 9 years 7 months and 13 years 6 months, with respective ranges of 8 years 7 months to 10 years 2 months and 12 years 8 months to 14 years 4 months. Each subject received five toffees and five balloons for the cooperation in the experiment.

Results. The mean error term decreased from fourth standard (M = 4.16) to eighth standard (M = 1.92) in prediction of performance in puzzle contest, F (1, 46) = 4.72, $p < .05$. However, the error terms were nearly the same for the fourth (M = 1.77) and eight (M = 1.37) standard children, F (1, 46) = 0.94 in prediction of performance in singing contest.

In general, predictions of performance in puzzle contest (M = 3.03) had bigger error term than those in singing contest (M = 1.57), F (1, 92) = 7.06, $p < .01$. Such a task difference was present with fourth standard children, F (1, 46) = 6.25, $p < .01$, but not with eight standard children, F (1, 46) = 0.98.

Experiment 2

Method. The stimuli, design, and procedure were exactly the same as in the singing contest condition of Experiment 1. The only notable change was that the subject was urged to use information about both the motivation and ability of stimulus student, and the experimenter checked all the five practice examples of the subject to ensure that he or she indeed used both types of information.

The subjects were 144 children. The kindergarten children were from the Thumbelina Nursery and Kindergarten School, Narayanpura, Ahmedabad, India. Children of Standards II, IV, VI, VIII, and XI were from the same population as in Experiment 1. There were 12 boys and 12 girls in each age group. Mean ages for the six groups of children were 4 years 6 months, 6 years 6 months, 8 years 7 months, 10 years 8 months, 12 years 8 months, and 15 years 7 months with respective ranges of 4 years 2 months to 5 years 4 months, 5 years 11 months to 7 years 3 months, 7 years 10 months to 9 years 4 months, 9 years 11 months to 11 years 8 months, 12 years 3 months to 13 years 7 months, and 14 years 9 months to 17 years 2 months.

Results. The mean error term for kindergarten, second, fourth, sixth, eighth, and eleventh standard children were 0.67, 0.97, 0.89, 1.01, 0.99, and 0.42 in order. There was no effect of age of subjects, F (5, 138) = 1.48, and so the null result of the singing contest condition of Experiment 1 was precisely replicated. This shows that response variability does not decrease over increasing age of children at least in prediction of performance in singing contest.

It should be noted that the magnitude of error term is uniformly lower in Experiment 2 than in other experiments by Srivastava and Singh (1986a, 1986b). Perhaps emphasis on the usage of both cues during practice session enabled the subjects to pay careful attention to the given information over all three trials of judgment.

Conclusions and Implication

Findings from the analyses of error variance from individual subject analyses of the four experiments on prediction of performance from information about motivation and ability of stimulus persons allow two conclusions. First, the mean error term in the individual child analyses sometimes decreases (Gupta & Singh, 1981; Srivastava & Singh, 1986b, Puzzle contest), sometimes increases (Srivastava & Singh, 1986a), and sometimes remains the same (Srivastava & Singh, 1986b, Singing contest) over increasing age of the subjects. Thus, the generality of the age-trend in response variability found by Anderson and Butzin (1978) is severely restricted, and Surber's (1984) description of a decrease in error term with an increase in age of children as a "well-known problem in developmental research" (p. 242) needs to be corrected.

Second, error variance in individual child analyses appears to be better predictable from the nature of cues in the integration task (Srivastava & Singh, 1986a) as well as nature of judgment (Srivastava & Singh, 1986b) than from the age of subjects. The result that the four-cue task invokes less response variability than do the two-cue tasks suggests that the error variance depends upon what enter as inputs in the integration task. The difference in error term of even two-cue tasks, for example, Mother x Neighbor and Teacher$_1$ x Teacher$_2$

designs (Srivastava & Singh, 1986a) and of puzzle and singing tasks (Srivastava & Singh, 1986b) further suggest that response variability may be attributed to the different meanings and importance of inputs for the judgmental tasks, and not necessarily to variable integrational capacity over ages. Future research should, therefore, take up the issue of association of nature of inputs and dimension of judgment with response variability of children of different ages.

One methodological implication of the present results for developmental research deserves mention. In research on integrational capacity of children, response variability along with the number of significant main effects over ages must as a rule be checked. Simply assuming that error variance remains constant or steadily decreases over increasing age may at time lead to erroneous conclusions about cognitive capacity of children.

References

Anderson, N.H. (1980). Information integration theory in developmental psychology. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), Information integration by children (pp. 1-45). Hillsdale, N.J.: Erlbaum.

Anderson, N.H. (1982). Methods of information integration theory. New York: Academic Press.

Anderson, N.H., & Butzin, C.A. (1978). Integration theory applied to children's judgments of equity. Developmental Psychology, 14, 593-606.

Gupta, M., & Singh, R. (1981). An integration theoretical analysis of cultural and developmental differences in attribution of performance. Developmental Psychology, 17, 816-825.

Kun, A., Parsons, J., & Ruble, D.N. (1974). Development of integration processes using ability and effort information to predict outcome. Developmental Psychology, 10, 721-732.

Leon, M. (1982). Rules in children's moral judgments: Integration of intent, damage, and rationale information. Developmental Psychology, 18, 835-842.

Miller, P.H. (1982). Children's and adults' integration of information about noise and interest levels in their judgments about learning. Journal of Experimental Child Psychology, 33, 536-546.

Srivastava, P., & Singh, R. (1986a). Prediction of exam performance by children: Evidence for utilization of four pieces of information. Developmental Psychology, Submitted August 1986.

Srivastava, P., & Singh, R. (1986b). Cultural, developmental, and task differences in prediction of performance: An information integration analysis. Child Development, Submitted September 1986.

Surber, C.F. (1984). Issues in using quantitative rating scales in developmental research. Psychological Bulletin, 95, 226-246.

Winer, B.J. (1971). Statistical principles in experimental design (2nd Ed.). New York: McGraw-Hill.

Author Note

This research was supported by Grant F1/146/80-RG from the Indian Council of Social Science Research, New Delhi. The author thanks Gaurang Bhatt, Neelang Desai, and Prabha Singh for their assistance in data analyses, and Ajay Shah and R. Usha for their cooperation in the completion of the project.

Correspondence concerning this article should be addressed to Ramadhar Singh, Organizational Behavior Area, Indian Institute of Management, Ahmedabad 380 015, Gujarat, India.

Table 1

Mean Error Term as a Function of Number of Cues in Judgmental Tasks and Age of Children

Number of cues in judgmental tasks	Age groups				
	KG	I	II	III	IV
Four-cue: $M \times N \times T_1 \times T_2$	1.00	0.57	0.64	0.54	1.36
Two-cue : $M \times N$	1.06	0.97	0.96	1.11	1.48
Two-cue : $T_1 \times T_2$	1.41	1.04	1.30	0.90	1.77

Note. The M, N, T_1, and T_2 denote opinion of Mother, Neighbor, $Teacher_1$, and $Teacher_2$, respectively. Data from Srivastava and Singh (1986a).