THE QUARTERLY NEWSLETTER OF THE
LABORATORY OF COMPARATIVE HUMAN COGNITION
William S. Hall and Michael Cole, Editors
Center for Human Information Processing
University of California, San Diego
July 1979, Volume 1, Number 3

The Effects of Context on Mother-Child Interaction: A Complex Issue*

Jay Belsky
The Pennsylvania State University

In a recent issue of this newsletter, Graves and Glick (1978) questioned a basic assumption of investigators studying mother-child interaction; namely, that maternal behavior is consistent across contexts. More specifically, they cautioned against generalizing results from one situation to another upon discovering that mothers were systematically more likely to interact with their young children when knowingly observed in a university laboratory than when unobtrusively watched during the same laboratory session. As a result of these findings, these researchers were led to conclude that "all analyses of mother-child interaction must attend to contextual variations that might be affecting the nature of the interaction displayed" (p. 45).

Beyond Group Means

In his recent critique of developmental psychology, McCall (1977) reiterated a basic point that has continually eluded developmentalists in their data analyses; that is, that the stability/instability of mean performance scores is potentially independent of the consistency/inconsistency of individual differences. Statistically this implies, of course, that the correlation coefficient is independent of the means of the two distributions entering into its calculations. From the standpoint of contextual effects, then, the consistency of individual differences across obtrusive-unobtrusive observations is potentially independent of mean differences between these contexts.

The relevance, as well as significance, of this point to the discussion of contextual effects on mother-child interaction is most dramatically evidenced by considering the possibility that the six mothers observed obtrusively and unobtrusively in the Graves and Glick investigation maintained their relative rankings across context (as evidenced by high correlations between identical variables). This would indicate that despite the dramatic effect that context exerted upon mothers' average performance, knowledge of being observed did *not* in fact serve to distort maternal behavior, but rather simply inflated it. Generalization across contexts would therefore be permissible as long as this inflation factor was accounted for. Further, if maternal behavior were to be correlated with some criterion, say the child's communication skill, for purposes of determining mother's influence upon this realm of functioning, then the data from the obtrusive observations would yield virtually identical correlations as that from the unobtrusive observations. The point to be made is simply that mean differences between observations made in different contexts can only be considered to demonstrate the distortive influence of some contextual factor, and thereby raise questions concerning the generalizability of the data gathered, when coupled with low correlations across contexts. So long as correlations between identical behaviors remain high across situations, contextual variations can only be considered to function as a microscope—magnifying the frequency of behavior rates while preserving the relationship between individuals.

As it turns out, Graves and Glick did evaluate the stability of individual differences; considerations of space simply kept them from reporting their findings. Since they observed that "correlations across the two conditions (obtrusive and unobtrusive) were low" (Graves, personal communication), it would seem, on the basis of the above reasoning, that context did indeed distort maternal behavior in their study.

A Model of Contextual Influences

The interface of individual differences and group means, as it pertains to the issue of contextual effects on mother-child interaction, can be illustrated best by a model of the hypothetical effects that context can exert on behavior (see Table 1). In Cell I of Table I, the previously discussed inflationary effect of context is illustrated; while mean scores differ significantly across contexts, individual differences remain fairly stable. In Cell II we see the most distortive effect that context can exert, indeed the one discerned by Graves and Glick. Contextual variation not only results in large group differences, but also produces inconsistent individual per-

* The author would like to thank Mary Kay Goode, Ann Walker, and, most especially, Fran Deutsch for their helpful comments on earlier drafts of this manuscript. Author's address: Division of Individual and Family Studies, The Pennsylvania State University, University Park, PA 16802.

formance. Cell III characterizes a "mirror effect"; in this situation behavior in one context is essentially identical to that observed in another; at the level of individuals and groups, performance is maintained. Finally, Cell IV represents the truly problematical situation, for upon discovering small differences between means, the investigator can easily be seduced into believing that contextual variation exerts no impact upon behavior. Examination of correlational data suggests just the opposite, however, since in this case (as in Cell II) the relative standing of individuals is highly unstable.

Table 1

Hypothetical Effects of Context on Behavior

MEAN COMPARISONS	Individual Differences: High Correlations	Individual Differences: Low Correlations
Large Differences	I	II
Small Differences	III	IV

Previous Research on Contextual Influences

Most research that has been concerned with the effect of context on parent-child interaction has failed to evidence appreciation of the complexity inherent in the study of contextual influences. In fact, upon reviewing research on the effect of situational structuring (e.g., free play versus problem solving tasks), physical setting (home versus lab), and observational obtrusiveness on mother-child interaction, only two investigations have been discovered that considered — or at least reported — the stability/instability of individual differences as well as group means. Both concern comparisons of mother-infant interaction at home and in the laboratory (see below).

The work of Patterson and Reid (1970) and Zegoib and Forehand (1975) considered only mean differences in behavior across contexts. Similarly, the only research in this area concerning the differential influence of obtrusiveness on the behavior of mothers from different social classes also focused exclusively upon mean scores (Randall, 1975). The findings from this large and well designed investigation support the speculations of Graves and Glick [see also Sroufe (1970) and Tulkin (1973)], and thus led Randall to conclude that "the reported superiority of the verbal environment of the middle class home may be partially a function of the art of observation itself" (p. 14). But the reported data do not speak conclusively to the issue of distortion — and thus cross-contextual generalization — as no information is provided concerning the consistency of individual differences across conditions of obtrusive and unobtrusive observation. The research on situational structuring (Smith, 1958; Streisguth & Bee, 1972; Zegoib & Forehand, 1975), as well as the early work on home-lab differences (O'Rourke, 1963; Shalock, 1956; Moustakas, Sigel & Shalock, 1956), suffers from the same conceptual/analytic weaknesses; once again, then, conclusions regarding the inappropriateness of generalizing data from highly structured to loosely structured contexts, or from the laboratory to the home, are unwarranted as only mean differences and not test-retest reliabilities are reported across contexts.

Fortunately, the more recent work in the area of home-lab differences justifies the drawing of such conclusions since both group means and individual differences are considered. In finding high correlations as well as few mean differences (Cell II) across identically and highly structured situations in the home and the lab (mothers required to sit on floor and play with their one-year olds with six specified toys), Peterson (1975) was able to conclude that, under such conditions, setting exerts minimal effect upon mother-infant interaction.

I drew somewhat different conclusions in a more recent investigation comparing mother-infant interaction at home and in the lab (Belsky, 1977). I wanted to evaluate the commonly held assumption that interaction observed in this locale is representative of, and thereby generalizable to, such interaction as it occurs under more everyday circumstances in the real world. So I created distinctly different situations at home and in the lab. Specifically, in the lab a free play situation was created, analagous to those employed in other studies, by directing mothers to pretend they were home with free time on their hands. In order to maintain a naturalistic, everyday context in the home, mothers were instructed to go about their regular household routines, pretending the observer was not present. While not permitting specification of the causes of any observed home-lab differences (i.e., setting or instructions), this design did enable me to determine whether one oft-employed laboratory situation, the free play context, elicits interaction representative of that occurring under more naturalistic conditions in the home as it is commonly assumed. In order to guarantee that any observed home-lab differences were actually a function of context effects (defined as the combined influence of setting plus instructions), rather than the general instability in mother-infant interaction from one observation to the next (a concern not sufficiently appreciated in studies of context effects), half the dyads were seen twice in a single locale, the home (H-H) or the lab (L-L), while the remaining 12 were seen once in each setting, with order counter-balanced (H-L and L-H).

Home-lab comparisons revealed, surprisingly, marked differences between contexts, with mothers more frequently attending to, speaking to, responding to, stimulating and praising their infants in the lab while

prohibiting and ignoring them more often at home. That these differences were the result of observation context and not simply a function of behavioral instability across observation sessions, was demonstrated by the absence of differences between sessions for mothers seen twice in a single setting (H-H and L-L groups). Across session correlations (test-retest reliabilities) on summary indices of maternal activity and responsivity, it is important to note, were (non-significantly) higher for these mothers seen twice in a single location than for those seen once in each locale. Given both these sets of data (means and correlations), I was able to conclude that it is unjustified to presume that maternal behavior observed in a free play lab context is representative of, and thereby generalizable to, that which transpires under more everyday circumstances at home. Like Graves and Glick, I cautioned against making such cross-contextual generalizations until consistency across contexts (in terms of individual differences and group means) can be demonstrated.

REFERENCES

Belsky, J. Mother-infant interaction at home and in the laboratory: A comparative study. Paper presented at the biennial meeting of the Society for Research in Child Development (SRCD). New Orleans, March, 1977.

Graves, Z., & Glick, J. The effect of context on mother-child interaction: A progress report. *The Quarterly Newsletter of the Institute for Comparative Human Development,* 1978, *2,* 41-46.

McCall, R. Challenges to a science of developmental psychology. *Child Development,* 1977, *48,* 333-344.

Moustakas, E. E., Sigel, I. E., & Shalock, H. D. An objective method for measurement and analysis of child-adult behavior interaction. *Child Development,* 1956, *21,* 109-134.

O'Rourke, J. Field and laboratory: The decision making behavior of family groups in two experimental groups. *Sociometry,* 1963, *26,* 422-435.

Patterson, G. R., & Reid, J. B. Reciprocity and coercion: Two facets of social systems. In C. Neuringer and J. Michael (Eds.), *Behavior Modification in Clinical Psychology,* New York: Appleton, Century, Crofts, 1970.

Peterson, M. The study of a methodology for observing mother-infant interaction patterns during the second year of life. Unpublished doctoral dissertation, Cornell University, 1975.

Randall, T. An analysis of observer influence on sex and social class differences in mother-infant interaction. Paper presented at biennial meeting of the Society for Research in Child Development, Denver, Colorado, April, 1975.

Shalock, H. Observation of mother-child interaction in the laboratory and in the home. *Dissertation Abstracts,* 1956, *16,* 707.

Smith, H. T. A comparison of interview and observation methods of maternal behavior. *Journal of Abnormal and Social Psychology,* 1958, *57,* 278-282.

Sroufe, L. A. A methodological and philosophical critique of intervention oriented research. *Developmental Psychology,* 1970, *2,* 140-145.

Streisguth, A. P., & Bee, H. L. Mother-child interactions and cognitive development in children. In W. W. Hartup (Ed.), *The young child: Reviews of research,* Washington, D. C.: NAEYC, 1972.

Tulkin, S. An analysis of the concept of cultural deprivation. *Developmental Psychology,* 1972, *6,* 326-339.

Zegoib, L., & Forehand, S. Maternal interactive behavior as a function of race, socioeconomic status and sex of the child. *Child Development,* 1975, *46,* 564-568.

Response from Graves:

When asked to submit a summarization of my research for the *ICHD Newsletter* I was faced with the task of condensing a 150-page research report into an article of 10 short pages. In addition, the data needed to be presented, and the text reworded, so that the information would be comprehensible both to professionals in other fields and to general readers. During the summarization process, it was suggested that I take the scores for the individual mother-child dyads, which had been carefully compared across the two conditions, and combine them into one mean score for publication purposes. Although the base measures for the scores had differed for each individual mother-child pair, correlations across the two conditions for each variable were consistently low, and changes were in the same direction, for *all* the dyads. The scores were combined simply to indicate to the general reader, in a short-hand fashion, the basic "gist" of the findings.

I had assumed that it was clear from the way the text of the article was worded that the data for individual subject pairs had been analyzed, and that gross differences had been consistent across all subjects. In retrospect, I can see that an additional footnote would have been useful to indicate that data had been condensed for simplicity and due to space constraints. I'm sorry if my failure to do so generated confusion.

As a matter of fact, I happen to feel strongly, as does Belsky, about the distortion that can often be inherent in "averaged" data. I am very wary when I see such data reported and it was not until the point of actually submitting this shortened article that I even bothered to calculate means across all subjects.

I think Belsky's point about the difficulties in generalizing findings from lab to home is well-taken. People tend to produce different behavioral displays in different settings and for different audiences. Ideally I would also have liked to conduct the Graves and Glick study in the home environment, however, it would have

been difficult to videotape mother-child interaction in the subjects' home without their being aware of the process. My claim, however, was not that a sample of "naturalistic" behavior was being captured in either condition of our study, but that the nature of the mother-child interaction changes, at least for this white middle-class sample, when the presence of an observer with recording equipment is known.

I have no doubt that there is an additional effect of setting (i.e., home v. lab) on the nature of mother-child interaction, but I would question whether or not Belsky's sample of home data is respresentative of what might occur under "everyday circumstances," if there were no observer there whatsoever. Although the mothers are instructed to ignore the presence of the observer, I would doubt that there are many instances where this truly happens.

The Competence/Incompetence Paradox in the Education of Minority Culture Children*

Ronald Gallimore
Center for the Health Sciences
Department of Psychiatry and Biobehavioral Sciences
University of California, Los Angeles

Kathryn Hu-Pei Au
Kamehameha Early Education Program
and University of Illinois, Urbana-Champaign

A paradox is evident in many efforts to improve the educational opportunities available to minority culture children. While the children appear well-adjusted and entirely competent in their home environments, they often exhibit inappropriate behavior in the classroom and are slow to learn academic skills and content. It has occurred to some researchers and educators involved in the development of curricula for these children that their performance in school could be greatly improved if the abilities shown in the home environment could somehow be transferred to the classroom. In practice, this effort has proved more difficult than might be supposed.

The purpose of this paper is to address the problem of the home competence/school incompetence paradox as it applies to minority culture children. While this discussion will be based on our experiences in working with Hawaiian children at the Kamehameha Early Education Program (KEEP) in urban and rural Oahu, we believe this research suggests ways in which the educational achievement of other minority culture groups might also be improved.

The KEEP Reading Program

KEEP was begun in 1971 to solve through research and development the problem of teaching Hawaiian children who are educationally at risk to read at or above average levels (Tharp & Gallimore, in press). The conditions under which this effort was conducted must be considered nearly optimal: funding was stable, researchers had full control over the operations of a research and demonstration school, and professionals and scientists from a variety of disciplines (psychology, education, anthropology, and linguistics) participated. Still, the initial outcome was discouraging; during the first 2½ years, students in the KEEP research school read no better at the end of first grade than public school counterparts who generally score at the second stanine on standardized tests.

Beginning in 1976 and continuing through Spring 1978, a new reading program was employed; in its current form this curriculum is known as the Kamehameha Reading Objective System (KROS; Crowell, 1978; Tharp, in preparation).

Administrations of standardized reading tests indicated improved performance for the KEEP children taught with KROS while there was essentially no change in the scores of comparable public school controls during the same period. Indeed, the KEEP students on average, read near the national norm for their respective grade levels (Gallimore, Tharp, & Sloat, in preparation [a]).

Quite naturally, a number of different explanations for the success of the KEEP reading program have been suggested. We will argue that the KEEP reading program or KROS is successful, in part, because it resolves the competence/incompetence paradox by encouraging the more consistent application of cognitive strategies already in the children's repertoire.

Cognitive Strategies and the School Performance of Minority Culture Children

Brown suggests it is well established that early in their school careers, disadvantaged children have "difficulty generating aids, mnemonics, research strategies, etc., to enhance deliberate learning" (in press, ms. p. 92). Without these cognitive strategies a child is greatly handicapped in the ordinary classroom where tasks are presented in the absence of a meaningful framework; this is sometimes called a decontextualized instructional style. If children are accustomed in the home environ-

*This research was supported by the Kamehameha Schools and the Estate of Princess Bernice Pauahi Bishop. Assistance was also provided by the Sociobehavioral Research Group, mental Retardation Research Center, University of California, Los Angeles. Requests for reprints should be addressed to Ronald Gallimore, Department of Psychiatry and Biobehavioral Sciences, UCLA, Los Angeles, California 90024. Copies of KEEP Technical Reports and work in preparation may be obtained from Editor, Technical Reports, KEEP, 1850 Makuakane Street, Honolulu, Hawaii 96817.

Our appreciation is due Dr. Roland Tharp and Dr. Andrea Zetlin for critical reading of the manuscript.

ment to using cognitive strategies for which there are external cues, they may not be prepared to generate and selectively apply internally mediated aids. There are numerous examples of external aids, but for young children many are likely to be embedded in situations in which they interact with adult socialization agents (Wertsch, 1978).

There seems to be general agreement that children and adults who employ self-generated cognitive strategies perform better on school-type tasks than those who do not. Presumably because of the greater continuity between home and school, middle-class children are much more likely to use "school-efficient" internally mediated cognitive strategies than culturally and socially disadvantaged children (Brown, in press). The ready use of these strategies allows for more rapid adaptation to the school's learning style in which content is likely to be unrelated to daily life and initially meaningless to the child.

If the failure to use self-generated cognitive strategies accounts for the poor school performance of disadvantaged minority culture children, then it is important to specify exactly how they are involved. Some researchers have assumed that disadvantaged children lack certain school-relevant cognitive strategies. According to this "cognitive deficit" hypothesis, we would expect to find uniformaly low performance on all school-type tasks requiring the use of cognitive strategies.

The results obtained at KEEP, however, do not fit the pattern predicted by this assumption. We can summarize the KEEP results by stating that we do not observe uniformly low performance on all school-type tasks; what we see instead is a widespread *inconsistency* in performance across tasks and settings. This finding of inconsistency suggests that the children possess many of the same cognitive strategies as more school-successful middle-class children; the reason their school performance is so much poorer is that they apply the cognitive strategies much less consistently than their middle-class peers.

We further speculate that inconsistent application results from discontinuities between the school and home environments. A middle-class child is more likely to identify correctly the type of cognitive strategy called for in a given school task because of the similarity of elements, e.g., task and process variables, to those in the home. Greater differences between the school and home contexts for disadvantaged children make it much less likely that they will recognize the task as one which calls for the use of a certain cognitive strategy, although that same strategy might be readily used in home situations.

In short, we argue not only that cognitive strategies are implicated in the success of the KROS reading program, but also that the program specifically encourages more consistent use of classroom-useful cognitive strategies. In some cases these strategies may be developed through the KROS program, but the program also appears to function to make explicit to the children which cognitive strategies are required in particular situations.

Our argument is supported by data from (1) standardized testing, (2) experiments, and (3) formal and informal observations.

Standardized testing. During the time that a phonics-emphasis reading program was in effect at KEEP, the children did not learn to read any better than public school comparison groups. Yet during the years of phonics-instruction their scores on general cognitive and verbal ability tests such as the WPPSI and WISC-R were average or better (Gallimore, Tharp, & Sloat, in preparation [b]). After the KROS curriculum was installed, the children achieved rapid and impressive gains on standardized reading tests, relative to controls.

Two points relevant to our thesis will be emphasized here. First, it is evident that the KEEP students do not exhibit uniformly low performance on all school-type tasks; rather, after first grade they perform, as a group, at average levels on standard IQ tests, which incorporate many school-type tasks. Second, the relatively short time—kindergarten and first grade—required for the children to obtain average IQ scores would not allow for the development of basic cognitive strategies, if such capacities were entirely absent to begin with. The new reading program did not remediate deficits in fundamental processes; rather its effects appear to be on variables not tapped by the general ability tests.

Experimental evidence. Experimental data indicate that adults can easily elicit use of cognitive strategies by KEEP students. In one study, KEEP kindergarteners showed significantly better long term recall of shape names when they were prompted by an experimenter to associate the shape name with a common-place object, e.g., circle-plate; octagon-stop sign, etc. (Gallimore, Lam, Speidel, & Tharp, 1977). Other students learned a labeling strategy which they did not generalize to similar stimuli until given an augmented explicit prompt (Speidel, Hao, & Gallimore, 1976). Finally, a receptive adult can elicit impressive linguistic performances from primary-grade children, performances which reflect substantial cognitive complexity (Watson-Gegeo & Boggs, 1977).

Observational Evidence. Informal observations at the KEEP research school point to the involvement of cognitive strategies in the improved reading performances. These observations suggest that before the KROS curriculum was introduced, KEEP students generally used inefficient and lower-level learning strategies, in particular when faced with the type of episodic tasks which predominated in the previous phonics-oriented reading program. Among the observations were these: (1) when given a slightly new or altered task, KEEP students often failed to use skills/knowledge they had been observed to use on similar tasks; (2) unless directly prompted they usually did not relate personal knowledge and exper-

iences to school tasks; (3) they were likely to adopt a passive rather than active learning role; (4) guessing and other rote learning strategies were frequently observed; and (5) each problem was typically approached as a new and different task rather than as an instance of a class to which an already mastered solution might be applied.

After the new KROS reading curriculum was installed, the learning efficiency of the children appeared much improved. Observations suggest that the children made more use of deliberate strategies for learning; they became active and involved; answers were altered and changed. A child might make a response which was not accepted, and minutes later reintroduce the topic and offer an alternative response. Taken together, these observations suggest that Hawaiian children can be taught more efficient learning or cognitive strategies, an achievement which seems unlikely if generalized deficits were the main problem. In this respect the observations are consistent with previous ethnographic and behavioral studies which indicated that Hawaiian children are competent learners in the home environment (Gallimore, Boggs, & Jordan, 1974; Gallimore & Howard, 1968). Evidently KROS provided a bridging experience which encouraged and taught the children to perform at school at a level consistent with their home performance.

KEEP Reading Lessons and Cognitive Training

The KROS differs from the previous phonics program in its curricular emphasis. At present, more attention is given to direct instruction of the understanding of what is read, while previously the focus was on the learning of phoneme-grapheme correspondences, or phonics. The KROS attempts to strike a balance, by directly teaching phonics *and* comprehension simultaneously, an approach which has been recently described as urgently needing a thorough evaluation if progress is to be made in reading theory and practice (Resnick, in press).

What are these small group lessons like? Each KEEP class of 25-30 is divided into from 4 to 6 groups for reading instruction on the basis of criterion-referenced test scores. Within these more or less homogeneous groups of 3 to 8 children, the teacher is able to interact with each child more frequently and to informally assess his or her progress more accurately. Each group meets with the teacher for 20 minutes per day; a fixed amount of daily instruction is thus guaranteed for each child, although additional individual help may be given as time permits. At other work stations, sight vocabulary, language awareness, listening skills, and word attack skills are taught through individualized activities.

The teacher-led lessons are almost always based on a story from the children's basal text and consist primarily of interchanges in which the teacher asks questions which the children are expected to answer, interspersed with segments of silent reading. On first inspection the most impressive feature of these lessons is the constant give and take between teacher and child; the teacher continually asks questions to which the children respond. The types of questions asked by the teacher are intended to develop the children's proficiency in the hierarchy of comprehension skills specified in the reading curriculum (see Crowell, 1978). Upon closer inspection these reading lessons are shown to have a consistent structure. This conclusion is based at present on only a small sample of teachers whose lessons were thought to exemplify the type of instruction desired in the KEEP reading program. The analysis was developed through the study of videotapes of three reading lessons, each taught by a different teacher, one lesson in a first-grade class at KEEP, and two with KEEP second graders (Au, in press).

On the basis of transcripts it was found that each lesson could be divided into topically defined interchanges. For example, the key questions marking the interchanges at the beginning of the first-grade lesson were:

(1) Why does this story seem familiar?
(2) What does "make music" mean?
(3) What do you think Jasper will do to make music?

The topical interchanges fell into three categories: *E* or experience, *T* or task, and *R* or relationship. An *E* or experience interchange is one in which the teacher has the children discuss experiences or knowledge they have which are related in some way to the story. For example, at the beginning of the first-grade lesson the teacher has the children talk about what the phrase "make music" means to them. After this first section of the lesson, the teacher then has the children read silently short parts of the story, usually a page or two, asking them questions about the content of the story after each section read; these are the *T* sequences. In the final category of interchanges, the *R* interchanges, the teacher attempts to draw relationships for the children between the content of the story discussed in the *T* interchanges and their outside experience and knowledge. Thus the *R* interchanges provide for the integration of information contained in the *E* and *T* interchanges. Examples of *E, T,* and *R* interchanges are presented in Au (in press).

The three types of interchanges appear to differ in terms of what is needed to produce an acceptable response; the sources of information and task requirements related to each are different. In *E* interchanges the underlying form of question is, "What information obtained prior to this lesson do you have about the subject of the story?" In *T* interchanges it is, "What is the information given by the text?", whereas in *R* interchanges it is, "What interpretation can be given to this story as a result of combining these two types of information?"

The *E, T,* and *R* categories and the set of structural rules for the lessons seem to reflect the teachers' efforts to give the children systematic practice in the application of the cognitive strategies related to understanding a written story. The three teachers are skillful questioners,

particularly adept at leading children to the correct answers, rather than telling them the answers directly! This tactic seems to be an important aspect of the lessons. First, answering gives a child practice in producing the right kind of information at the right time. Second, after hearing a child's response, the teacher can determine which steps in the process are easy for each individual and which are more difficult, so that ensuing questions can be adjusted to the proper level.

As children gain more experience, it seems reasonable to expect that they will begin to apply consistently on their own the same cognitive strategies that the teacher encouraged in the reading lessons. It could thus be hypothesized that a child who has interacted frequently with a teacher in these consistently structured lessons would show better reading comprehension than a child who has not had the same experience. In fact, this is what the KEEP reading test results indicate may have happened (Gallimore, Tharp, & Sloat, in preparation [a]; Tharp, in preparation).

For Hawaiian children, who often do not succeed in school, *E* interchanges may be of particular importance in learning to read, because they can serve as links between the home and school environments. The teacher's use of *E* interchanges thus signals to children that certain cognitive strategies already in their possession may be relevant in an otherwise unfamiliar school situation, the reading lesson. By beginning the lesson at the level of the child's own experiences (*E* interchanges), the teacher increases the probability that existing cognitive strategies will be applied. When unfamiliar story content is introduced (*T* interchanges), it is presented within the context of content familiar to the child. Finally, by making explicit to children the relationships between their own knowledge and information in the text (*R* interchanges), the continued steady application of cognitive strategies throughout the lesson is encouraged.

What Strategies Does the Program Affect?

Why did KEEP fail to teach reading to the Hawaiian children before the new program was introduced? Why did motivated, well trained, and dedicated teachers using a solid, dependable curriculum not teach these competent children to read? After all, most children learn to read, and are taught to do so with a wide variety of approaches. It seems safe to assume that one of the commonalities among these highly varied sets of circumstances is that they do not prevent children from learning to read. This notion, it will be seen, is analogous to Scarr-Salapatek's (1976) idea that development is supported by a wide variety of seemingly very different environments; another way to think of the problem is to consider those few circumstances under which it does not proceed apace (Flavell, 1977). This point of view may seem somewhat implausible until we consider the following fact: two long-term research projects obtained convincing evidence that Hawaiian children should learn how to read (Gallimore, Boggs, & Jordan, 1974; Tharp & Gallimore, in press), including experimental, observational, psychometric, and ethnographic data. It may be that we have, in our failure, made an important discovery — we have succeeded in identifying one of the ways to *prevent* Hawaiian children from learning to read. How does the process of prevention work? We might hypothesize, in the case of reading, that it operates by interfering with the naturally developing strategies of the child.

Now we are able to turn to the long-deferred issue of defining more precisely what these strategies are. We hypothesize that the original phonics curriculum interfered with the application and natural development of *top-down processing strategies* through an overemphasis on *bottom-up processing strategies.* Top-down processing begins with already existing language and knowledge; bottom-up processing begins with recognition of letters and words and proceeds to comprehension; sometimes this distinction has been described as episodic versus semantic learning or processing.

In the mind of the competent reader, bottom-up and top-down processing occur simultaneously. Information continually flows from top to bottom and from bottom to top so that the results of analysis at any level may serve to facilitate further analysis at every other level (Adams, in press; Rumelhart, 1976). Too heavy reliance on top-down processing will mean existing knowledge and understanding may obscure the content of the text so the reader fails to comprehend the author's message and ignores visual information necessary for the accurate recognition of words. Conversely, use of strict bottom-up processing may lead to the identification of individual words without an understanding of the meaning of the text.

Indirect evidence that the phonics program discouraged use of top-down strategies by KEEP students was provided by a study of oral reading errors (Au, 1976; 1977). Fifteen second-graders were identified as good and poor readers on the basis of standardized reading achievement scores. All children read the same stories; using tapes of the performances, judges categorized errors according to whether they reflected use of context, visual-phonic information, both, or neither. Among the conclusions, Au noted that the KEEP students, who had been taught a bottom-up (phonics) approach, appear to be less proficient in use of context than children who were subjects in similar studies in both the mainland U.S. and New Zealand. It also appears that the phonics-instructed KEEP students did not approach reading as a language task; they did not use their fluency in language to aid them in reading, and did not have the idea that trying to solve a problem in reading can be approached by testing linguistic hypotheses.

Au (1976) does not present data for error patterns with comprehension-emphasis instruction, so it is not

certain that the conclusions of her study can be used to support the present argument. However, these findings are consistent with the hypothesis that Hawaiian children, taught with a phonics or bottom-up approach, do not make use of existing skills — e.g., language fluency — to facilitate top-down processing. Instead, the children in Au's study relied on bottom-up strategies, in this case visual-phonic, rather than top-down strategies such as context.

In a strict phonics approach the child may learn to focus exclusively on the phonemic-graphemic and visual-phonic cues while systematically ignoring the semantic ones. This is a major criticism of such programs (Adams, in press), one which may be particularly apt in the case of minority culture children. While other populations of children apparently do learn to read in programs where there is a heavy emphasis on phonics, the KEEP students did not. It is evident that there are interactions between program effects and the characteristics of different groups of young readers; specifically, the top-down processing strategies of the KEEP students may be more sensitive to disruption than those of middle-class children because of the greater discontinuity between home and school, in terms of physical features of the environment, styles of interaction and learning and types of content. If there is a pre-existing bias for inconsistent application of these strategies due to the nature of the school setting and the child's resulting uncertainty about its demands, a strict bottom-up reading curriculum might well suppress use and development of top-down reading strategies. But if the child's natural impulse is to make sense out of all but the most completely nonsensical situations, comprehension (top-down) strategies will emerge as soon as they are encouraged and taught. We hypothesize the KEEP children readily began to use such strategies once the reading program was altered. Their reading scores then improved dramatically, to a level commensurate with their scores on the general ability tests.

We believe that including top-down approaches in instruction may increase the likelihood of minority children applying existing knowledge, linguistic skills, and strategies to unfamiliar classroom situations. Exclusively bottom-up approaches may cause the children to view culturally unfamiliar tasks as episodic, rote, decontextualized activities unrelated to previous experience. In contrast, majority culture children may find school work more familiar and see more rapidly the relevance of ideas and strategies learned at home, and thus develop and use appropriate top-down strategies even when taught with a phonics curriculum. We do not know whether the progress of majority culture students would also be improved by a better balance between bottom-up and top-down strategies in instruction. However, provision of ETR interchanges increased listening and reading comprehension skills of moderately retarded adolescents (TMRs), a conclusion based on pre- and post-testing, weekly criterion tests, and transcript analysis (Zetlin & Gallimore, 1979). One explanation of these results is application to written text of top-down processing which retarded students are rarely encouraged to use; most special curricula feature low-level tasks and do not promote higher-order cognitive processes on the assumption that TMRs learn best by rote and repetition (Levine, Zetlin, & Langness, 1979; Winschel & Enscher, 1978).

According to Resnick (in press) there are no fully documented accounts of reading programs which simultaneously feature top-down and bottom-up processing, with any populations including middle-class majority culture groups. Research in progress at KEEP is designed to test the hypothesis that simultaneously fostering top-down and bottom-up processing is a key factor in the success of Hawaiian children taught with the KROS program. The outcome of this effort will also bear on the argument that supression of top-down processing strategies produces the competence/incompetence paradox so often observed in minority culture children.

REFERENCES

Adams, M.J. Failures to comprehend and levels of processing in reading. In R.J. Spiro, B.C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension.* Hillsdale, N.J.: Erlbaum, in press.

Au, K.H. An analysis of oral reading errors and its implications for reading instruction (Tech. Report #50). Honolulu: The Kamehameha Schools. The Kamehameha Early Education Program, 1976.

Au, K.H. Analyzing oral reading errors to improve instruction. *Reading Teacher,* 1977, *31*(1), 46-49.

Au, K.H. The role of the adult in the development of metacognitive processes related to reading. Unpublished manuscript, 1978.

Au, K.H. Using the E-T-R method with minority children. *Reading Teacher,* in press.

Brown, A.L. Knowing when, where and how to remember: A problem of metacognition. In R. Glasser (Ed.), *Advances in instructional psychology.* Hillsdale, N.J.: Erlbaum, in press.

Crowell, D.C. Kamehameha Reading Objective System. Honolulu: Kamehameha Early Education Program, 1978.

Flavell, J.H. *Cognitive development.* Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Gallimore, R., Boggs, J.W., & Jordan, C. *Culture, behavior and education: A study of Hawaiian-Americans.* Beverly Hills: Sage, 1974.

Gallimore, R., & Howard, A. (Eds.) *Studies in a Hawaiian community: Na makamaka O Nanakuli.* Pacific Anthropological Records No. 1, Department of Anthropology, B.P. Bishop Museum, Honolulu, 1968.

Gallimore, R., Lam, D.J., Speidel, G.E., & Tharp, R.G. The effects of elaboration and rehearsal on

long-term retention of shape names by kindergarteners. *American Educational Research Journal,* 1977, *14*(4), 471-483.

Gallimore, R., Tharp, R.G., & Sloat, K.C.M. Analysis of achievement test results for the Kamehameha Early Education Project: (1972-1978). KEEP Technical Report, in preparation (a).

Gallimore, R., Tharp, R., & Sloat, K.C.M. Analysis of WPPSI and WISC-R scores for the Kamehameha Early Education Project: (1972-1978). KEEP Technical Report, in preparation (b).

Levine, H., Zetlin, A., & Langness, L. Everyday memory tasks in a school for the TMR learner. Paper presented at the 103rd Annual Meeting of the American Association on Mental Deficiency, Miami, Florida, May, 1979.

Resnick, L.B. Theory and practice in beginning reading. In L.B. Resnick & P.A. Weaver (Eds.), *Theory and practice of early reading.* Hillsdale, N.J.: Erlbaum, in press.

Rumelhart, D.E. Toward an interactive model of reading (Technical Report #56). Center of Human Information Processing, University of California, San Diego, 1976.

Scarr-Salapatek, S. An evolutionary perspective on infant intelligence: Species patterns and individual variations. In M. Lewis (Ed.), *The origin of intelligence: Infancy and early childhood.* New York: Plenum, 1976.

Speidel, G.E., Hao, R., & Gallimore, R. Production deficiency of labeling skills in a pre-reading letter discrimination task (Technical Report #37). Honolulu: The Kamehameha Schools. The Kamehameha Early Education Program, 1976.

Tharp, R., & Gallimore, R. The ecology of program research and development: A model of evaluation succession. In L.B. Sechrest (Ed.), *Evaluation Studies Review Annual* (Vol. 4). Beverly Hills: Sage, 1979, in press.

Tharp, R. The direct instruction of comprehension within an integrated reading curriculum: Description and results of the Kamehameha Early Education Program (KEEP). KEEP Technical Report, in preparation.

Watson-Gegeo, K.A., & Boggs, S.T. From verbal play to talk story: The role of routines in speech events among Hawaiian children. In S. Ervin-Tripp & C. Mitchell-Kernan (Eds.), *Child discourse.* New York: Academic Press, 1977.

Wertsch, J. V. Adult-child interaction and the roots of metacognition. *The Quarterly Newsletter of the Institute for Comparative Human Development.* 1978, *2*(1), 15-18.

Zetlin, A.G., & Gallimore, R. Can TMR's think? Paper read at Gatlinberg Conference on Research in Mental Retardation. Gulf Shores, Alabama, April, 1979.

Winschel, J.F., & Ensher, G.L. Curricular implications for the mentally retarded. *Education and training of the mentally retarded,* 1978, *13,* 131-138.

A Comparative Analysis of the Acquisition of Numeration: Studies from Papua New Guinea*

Geoffrey B. Saxe
The Graduate Center of the City University of New York

Cultural groups have invented numerational systems to represent numerical information, and the organization of these systems varies from one group to another. For instance, in the West, we use arbitrary signs for numerals and our numerational system has a base of ten. In contrast, in Papua New Guinea, many groups use body parts as numerical symbols and the bases of these counting systems vary widely. Some of these systems have no base structure at all.

Both Vygotsky (1962, 1978) and Piaget (1970) have offered formulations of the relation between the child's acquisition of sign systems (of which numeration is one) and the development of intelligence. In contrast to conventionalist interpretations of knowledge that reduce mathematical and logical cognition to conventional systems of symbolization, both Vygotsky and Piaget maintain that the origins of logical thought are not reducible to conventional sign systems.

Piaget's formulation of cognitive development focuses on universal forms of cognition and changes in the structure of these forms over the course of development. During infancy, the structure of cognition is limited to sensori-motor coordinations, or "practical operations." Over the course of development, the child achieves progressively more powerful levels of operational coordination, and these operations achieve representational forms. It is this developing coordination that enables the child to structure the numerational system into a sign system for number, and the numerational system thus becomes an extension of the child's intellectual operations.

Vygotsky, in contrast, was not concerned with the structure of cognition. Rather, Vygotsky was concerned with how the products of the historical evolution of knowledge come to be interiorized and to transform the process of intellectual functioning of the individual. For instance, Vygotsky maintained that speech and thought have different genetic roots. Speech undergoes a preintellectual phase in its early development and thought a

*Only with the cooperation, help, and good will of many people was this work possible. I am especially indebted to David Lancy, Principal Research Officer of the Department of Education in Papua New Guinea, to Virginia Guilford and Tom Moylan, two in-resident anthropologists in the Oksapmin area, and to the kind and generous people of the village groups who allowed me to visit and become friends. The research was supported by grants from The Education Research Unit of Papua New Guinea (#534), The National Institute of Education (G-78-0076), and Sigma Xi. Appreciation is also extended to Maryl Gearhart who read and commented on a previous draft of this manuscript.

pre-verbal phase. At some point in early childhood, these two functions meet and begin to influence one another's formation. Thought transforms pre-intellectual speech into a mediational system, and the mediational system, in turn, comes both to serve and help organize the child's developing intelligence. With respect to numeration, Vygotsky's formulation suggests that as thought transforms the number words of language into a mediational aid, the child achieves not only the ability to represent quantity, but inherits a legacy of knowledge produced in social history, and it is this historical knowledge that comes to guide the individual's solutions to mathematical and physical problems.

The purpose of this paper is to present a summary of my research on the child's acquisition of different numerational systems in Papua New Guinea. An underlying assumption that guided this research was that there are two ways in which cross-cultural analyses of numeration can contribute to our understanding of cognitive development. First, one type of cross-cultural analysis can reveal cultural universal and cultural specific processes in the acquisition of a numerational system. The aim would be to determine, in accordance with a Piagetian formulation, whether there are universal intellectual operations that are necessary preconditions for the acquisition of any numerational system and, if so, the way in which these operations interact with a particular numerational system to produce cultural specific patterns in the acquisition process. Second, another type of cross-cultural analysis can be used to determine whether the acquisition of a particular numerational system, once interiorized as a "tool" of the child's developing intelligence, leads to cultural specific differences in the child's elaboration of physical and mathematical concepts (e.g., the further development of arithmetic operations such as formal addition, subtraction, multiplication, and division). In accordance with a Vygotskian approach, such an analysis would contribute to our understanding of the functional relations between the socio-historical evolution of representational systems and the cognitive development of the individual. The research that I conducted among people living in Oksapmin, Melpa (Mount Hagen), and Ponam Island village communities of Papua New Guinea was devoted for the most part, to the first of these concerns, and it is this research that I will summarize in this report.

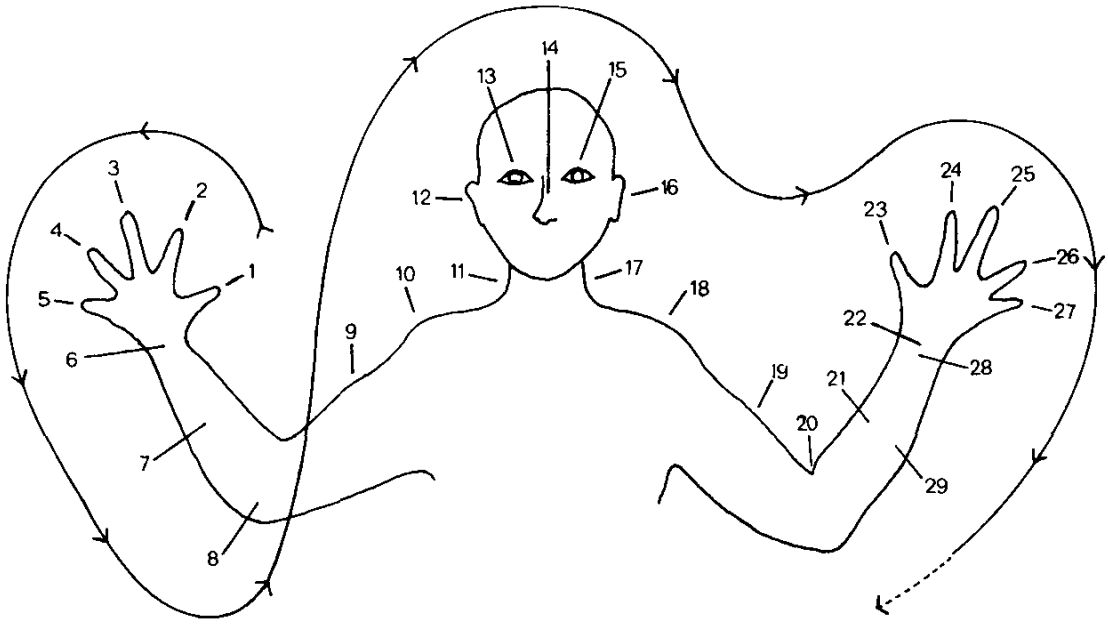
The Cultural Groups and Their Numerational Systems

Each of the village groups — Oksapmin, Melpa, and Ponam Island — differ from one another in many significant ways; these differences include the degree of assimilation of these communities to Western culture as well as the forms of the indigenous numeration systems that these groups use. Oksapmin and Melpa are both remote highland communities, although Oksapmin remains more removed from Western contact than the Melpa. People in the Oksapmin area are subsistence gardeners who cultivate vegetables (e.g., sweet potato and taro) and keep domesticated pigs; many of the members of the Oksapmin communities still wear indigenous dress (penis sheath and grass skirt). Among Melpa communities, signs of Westernization are visible to a greater degree. The local economy is in the throes of transition from a subsistence orientation to a commercial orientation; cash crops such as coffee beans are being cultivated by some, and sold to the Western cadres of the coffee industry. Ponam Island, a tiny island of the Admiralty chain off of Manus Island, has a very different physical environment than the highland groups. It is located amidst coral reefs, where fish are plentiful and are the major source of food. Other sources of food grow on the island itself, such as coconuts and tropical fruit, and there is weekly trade with people in neighboring villages on the main island for commodities unavailable on the tiny island itself. During World War II the island was evacuated by the United States and converted into an air base. At the end of World War II the Air Force left, but despite its remote location there has been continued contact and influence from the West.

The indigenous numerational systems among these communities differ. Both the Oksapmin and the Melpa systems use body parts to represent number, although the Melpa system has a base structure and the Oksapmin system does not. Ponam has a base-10 numerational system that is not of the body-part type. This system has many variations that differ with object-type enumerated (e.g., leaves, canoes, bananas). In addition, on Ponam, an indigenous birth order numerational system (or naming system) is used that differs in its organization from the core numerational system. In all groups, the Western numerational system is visible and is increasingly used. In the Oksapmin community, however, its use and influence is very minimal.

The Oksapmin body part numerational system has several variations. The one that I observed with the most frequency is organized such that 27 body parts are named in serial order between the thumb on the right hand and the little finger on the left hand (see Figure 1). As the individual ascends up the right side of the body in a count, body parts are named, and as an individual descends down the left side, body parts are named, but a prefix is named before each body part. If one needs to count further, one continues in a loop-like fashion to the wrist of the left hand and ascends back up the left side of the body.

The Melpa counting system also has several variations; the system that was most commonly reported to me was a base eight body part system. In this system, an individual counts the four fingers of each hand by twos, first bending down the small and ring fingers on one hand, then the two remaining fingers (middle finger and fore finger). The process is repeated on the other hand, and when completed, the two hands are brought



Oksapmin Body Part Counting System

together and the counter says "angika," which means 'one set.' If an individual needs to count further, the process is repeated again and the counter brings the hands together upon completion of another set and says "angika-kee," or 'two sets.' Unlike the Oksapmin system, the Melpa system, in principle, can make use of an infinite set of numerals since the system has a base structure.

On Ponam, the base-10 "core" numerational system is generally known by most adolescent and adult members of the community. The variations of this system that are linked to enumerating specific kinds of objects are becoming obsolete, however, and these variations are generally known only by elders on the island. Ponam's birth order numerational system is commonly known by both young and old members of the community. There are two variations of the birth order system that share a similar organization. The most commonly used system is one in which boys are given one of eight names depending upon their birth order relation to other boys in their family, and girls are given one of a different set of eight names depending upon their birth order relation to other girls in their family. The organization of this birth order system has the interesting feature that, although the names of same sexed children of a family by definition specify an order relation in age to one another, the names of opposite sexed children of a family do not by definition specify an order relation in age to one another. For instance, in a single family, *Pitatah* (name of second born girl) must always be older than *Pisiwah* (third born girl), but *Pitatah* (second born girl) may or may not be older than *Tol* (first born boy), since two or more girls of a family could be born before the first boy.

The Acquisition Process

The major theoretical hypothesis that guided the comparative research was that, consistent with both Piaget's and Vygotsky's formulations, children do not acquire the use of the numerational system of their culture as a ready-formed set of numerical meanings. Instead, the children, in effect, turn the numerical terms of the numerational system of their culture into a symbol system for number. Moreover, this transformation is accomplished by means of developing intellectual operations. In a recent study (Saxe, 1979b), I argued that these operations consist of the child's ability to coordinate cognitively a successive iteration of elements with their progressive summation. Young children can consider elements of a collection successively, and can consider a group of elements as a conglomerate, but they cannot, however, at once consider (conceptually coordinate) a successive iteration of elements and their progressive summation. This analysis suggests that young children may acquire facility with the number names of the counting system of their culture, learn them in their proper order, apply a single name to a group of objects, and, on occasion, even count accurately. Children should nonetheless evidence what Vygotsky termed a "pre-intellectual" (or pre-quantitative) use of their counting system. That is, children should not understand the numerical significance, or the "numerical referent" of their counting, since children cannot conceive of a progressive summation of individuals. As

children achieve this operational coordination, they achieve the intellectual preconditions for the structuration of their number words into a system for the notation of these operations.

In the West, evidence supporting this general formulation comes from various sources. For instance, in a previous study conducted with children from the United States (Saxe, 1977), I demonstrated that although young children often know the number names in order, they use these names in a "pre-intellectual" or "pre-quantitative" fashion when required to use counting to mediate numerical comparisons or numerical reproductions of collections of objects. For example, when making a numerical comparison between two collections, children may count but they do not base their numerical comparison on counting; instead, they base it on a non-numerical variable, such as the spatial extent of the collections. Shaeffer, Eggleston, and Scott (1974) report a similar phenomena. They found that although young children may or may not count an array accurately, they frequently do not understand that the last numeral recited can be used as a representation of the cardinal value of the collection. In the course of development, children achieve an understanding that the numerical terms can be used to represent numerical relations. Consider the following observation of a 4-year-old engaged with her counting.

> Annie (4 years; 2 months) who knows the number names by rote well beyond twenty, looks very focused as she accurately counts a collection of three small toy pigs. She then adds another and recounts the collection accurately, this time exclaiming "four," and recounts with excitement and interest. She continues this process, each time exclaiming the last value of her count, until the seventh pig, at which point she miscounts the collection to six. She looks a little surprised but, nonetheless, adds another pig and recounts the collection to eight. Now she seems confused and recounts the collection, again counting to eight. She then starts the entire process again, beginning with three.

Annie appears to be discovering (or confirming her knowledge of) the effect of the addition of a unit upon the sum of the collection. This observation suggests that the pre-intellectual phase in the use of numerals serves an important function for the development of the mediational system. It presents the child with the "cultural material" which comes to be organized by the child's developing operations.

Based on these data from the West, I hypothesized that in any culture children's transformation of the numerational system of their culture into their own vehicle for numerical representation should be identifiable as a progression from the use of pre-quantitative to quantitative forms of the use of the numerational system. Different numerational systems, however, should present different problems of quantification, or "stumbling blocks," for the child in the acquisition process, and therefore we should expect to observe cultural specific ways that the transition from the pre-quantitative to quantitative use of numerational systems is manifested in the course of development. In order to test this general formulation, two types of studies were conducted in each of the three Papua New Guinea cultural groups. The first type of study was designed to test the hypothesis that the progression from pre-quantitative to quantitative counting strategies is a universal process, regardless of the structure of the numerational system. These studies were conducted with each of the three cultural groups and were directly comparable to my previous studies in the United States. The second type of study was designed to reveal cultural specific features in the child's acquisition of numeration systems in the context of the general transformation from pre-quantitative to quantitative understanding of numeration. These studies were conducted with children from both Oksapmin and Ponam.

In the first type of study — the one that focused on the universal changes and the use of numeration — children were presented with tasks requiring numerical comparisons and numerical reproductions of collections of objects. In each cultural group, materials used in the tasks were indigenous to the communities. The tasks were conducted through the aid of a translator, and a back translation technique was used to help insure the accuracy of the translations. The Oksapmin children used the indigenous body part counting systems on these tasks; the children on Ponam Island and in Melpa, however, used the Western counting system (although they spoke in their native languages). The findings were the same across the different cultural groups and were consistent with those cited among children in the United States. The younger age groups tended to use pre-quantitative counting strategies in their comparisons and reproductions, whereas the older groups used quantitative counting strategies. Moreover, as in the West, the transition from pre-quantitative to quantitative counting strategies was interrelated with counting accuracy.

Among the Oksapmin and Ponam village populations, two additional sets of studies were conducted that focused on cultural specific features in the acquisition of numeration that might reveal some of the processes that underlie the transition from pre-quantitative to quantitative forms of numeration. One study focused on developmental changes that are specific to a body part numerational system (Oksapmin), the other to a birth order numerational system (Ponam).

Among the Oksapmin, I suspected that, as Schaeffer, Eggleston and Scott (1975) found in the United States, one of the causes of the young children's pre-quantitative use of their counting system was that they do not understand that the last numeral recited (or body part iterated) has a quantitative significance, that is, in the case of the Oksapmin child, that it represents a summation of all prior body parts iterated. If this were the case

then the Oksapmin child, even though capable of iterating body parts in their conventionally defined sequence, should tend to consider the physical or functional similarities between body parts in evaluating their equivalence or non-equivalence, rather than differences between the progressive summations implied by the body parts. Hence, I suspected that the young Oksapmin child should consider symmetrical body parts (e.g., left wrist, right wrist) as the same "numbers," despite the different ordinal positions of these body parts in an enumeration.

In order to test this hypothesis, I interviewed 48 Oksapmin children individually who ranged in age from 6 to 16 years and who had no more than four months of schooling. I asked each child to demonstrate how to count (if the child did not know, we did not continue the interview). I then told each child a story that was appropriate for Oksapmin village life (a long procedure of back translation and piloting was used to develop the stories) about a man who was counting sweet potatoes in his garden. The child was told that one day the man counted to a certain place on his body (e.g., his right wrist) while a sweeping gesture was made from the child's right thumb to the specified body part on the right side of the child's body (see Figure 1), and that another day the man counted to another place on his body while a sweeping gesture was made around the child's body from the right thumb to the specified place on the left side of the child's body (e.g., left wrist). On three occasions pairs of symmetrical body parts were indicated, and on three occasions pairs of asymmetrical body parts were indicated. After each "trial," the child was interviewed about which body part represented the bigger number of sweet potatoes. As predicted, the findings were consistent with the hypothesized trend. Young children tended to identify symmetrical body parts as equivalent numbers (although many responded correctly on the asymmetrical body part questions), whereas older children said that symmetrical body parts were different numbers and correctly identified the larger numerical value.

In order to corroborate the finding that Oksapmin children do not treat body parts as numerical symbols, despite the fact that they know the conventional body part counting system, I conducted an additional study with Oksapmin children. In the second study, children were told that, although in their village people count from the right to the left side of their bodies, in a village over the mountains people count from the left to the right side of their bodies. (This was demonstrated with sweeping gestures across the child's body, and we then checked the child's comprehension of this description.) The child was then told that an Oksapmin man was counting sweet potatoes in his garden and he counted to this number (a sweeping gesture was made from the left side of the child's body to a specified body part on the right side, e.g., the right shoulder), and that the other man from the other village was also counting sweet potatoes and counted to this place (a sweeping gesture was made from the right side of the child's body to the *same* body part on the same side of the child's body, i.e., the right shoulder again). The child was then asked whether the two men counted to the same number of sweet potatoes or whether they counted to a different number. As predicted, we found an age trend on this task which paralleled the previous findings. Young children claimed that the body part represented the same number, and older children claimed that the same body part represented different numbers and correctly identified the man who counted to the bigger number.

These two studies indicate that in the course of the Oksapmin children's acquisition of their numeration system, they experience a confusion about the symbolic reference of their body parts. Young children behave as if they do not understand that a body part refers to a progressive summation of body parts, or numerals; rather, they evaluate body parts with respect to their physical similarity. In contrast, older children have organized the conventionally defined sequence of body parts into a symbol system for number, such that each successive body part reflects a progressive summation of the earlier body parts.

The Ponam birth order numerational system differs from both the Oksapmin and the Western numerational system, and it therefore provided a new setting to study whether numerational systems present children with an organized set of numerical meanings or whether children actually structure the numerational system of their culture into a system of symbolization for numerical relations. Unlike notational counting systems, the structure of the birth order system is based upon a form of operational coordination different from a progressive summation of individuals; it is based upon transitive ordinal operations (see above description of the system). I hypothesized that, despite the fact that young children use and know the birth order names in their proper sequence, understanding the age relations between same sexed members of the family implied by the birth order names would be acquired with the emergence of the child's understanding of transitive relations (as defined by successful performance on a Piagetian seriation of sticks task), but understanding the indeterminateness of age relations of opposite sexed members of a family implied by the birth order names would necessarily follow reasoning about relations within sex in that it is based upon a more complex coordination involving two sets of transitive relations (i.e., coordinating two separate systems of asymmetrical relations with one another).

In order to test these hypotheses, I interviewed twenty-nine children at each of three different age levels: 8 to 10 year-olds (2 years of schooling); 14 to 15 year-olds (5 years of schooling); and 16 to 23 year-olds (at least 6 years of schooling). The study had three parts. During Phase I, children were checked for their rote knowledge of birth order names and administered

Piaget's seriation of sticks task to assess their understanding of order relations. During Phase II, each child was presented with two series of four cards aligned in parallel rows, one consisting of pictures of girls and the other of boys. Each card contained a number (1, 2, 3, or 4) and the appropriate birth order name of the child. Children were then presented with a ninth card of a mother and father and told to make a family of the eight children aligned in their proper order of birth in a single row. Once the child had completed a construction, the cards were returned to their initial configuration of two parallel rows; the child was then required to construct two additional orders of eight cards, such that all were properly named.

During Phase III, the cards were returned to their initial configuration of two parallel rows, and the child was interviewed about age relations between children of the same and opposite sex. For instance the child was asked who is older, *Tol* (first-born male) or *Pisiwa* (third-born female)? How do you know? Could *Pisiwa* ever be older than *Tol*? How do you know? If the child responded affirmatively, he or she was required to construct a family with the cards arranged demonstrating this possibility. Again, a back translation technique was used to help insure the accuracy of the translation of these procedures, and extensive piloting was completed with the procedure before it assumed its final form.

As predicted, analysis of children's performances revealed a clear developmental trend. The youngest age group had minor difficulty with Piaget's additive seriation task (50% of the children achieved an advanced performance, and 30% achieved a transitional performance), and the majority of children in this age group responded correctly across the Phase III questions about age relations within sex. In contrast, none of these children responded correctly to the questions about the indeterminateness of age relations of opposite sexed children. In the middle age group, all of the children solved the additive seriation problem and the within sex questions. Similar to the youngest age group, most of these children could not understand the indeterminateness of the across sex age relations, i.e., that a fourth-born boy might be older than a second-born girl. Finally, virtually all of the subjects of the oldest group responded with appropriate answers across all three phases of the tasks. As in the case of both the Oksapmin body part system and the Western numerational system, these findings indicate that the child's acquisition of a numerational system is dependent upon developing intellectual operations.

SUMMARY

Consistent with the constructivist formulations of both Piaget and Vygotsky, these studies on the acquisition of numeration indicate that children do not acquire the number terms of their culture as a set of ready-formed numerical meanings, but rather organize the conventionally defined numerical terms into a system of symbolization for numerical relations. In all groups studied, this process of organization was manifested as a general transformation from the use of pre-quantitative to quantitative counting strategies to mediate numerical comparisons and reproductions. Moreover, in both the Oksapmin and Ponam communities, it was also demonstrated that the characteristics of particular numerational systems present different problems of acquisition for children, and these differences lead to different patterns of acquisition.

The research reported in this paper did not address the way a sign system, such as numeration, once formed, alters the child's representation of quantity, and, as such, provides children of different cultural groups with different kinds of opportunities for elaborating mathematical and physical concepts. Such an approach is a particularly interesting means of analyzing one class of functional relations that obtain between the historical evolution of knowledge structures within a cultural group and the process of cognitive development of its individual members, and I am just beginning analysis of this type.

POSTSCRIPT

The comparative approach outlined in this paper makes use of a macro-variable — culture — to study the relations between a sign system, numeration, and the development of numerical cognition. I have also been making use of organismic factors to study the relations between numerational systems and numerical cognition. In particular, recent studies on the breakdown of the use of numeration following focal brain injury (Saxe, Note 1) and the atypical development of numerical cognition in children with various forms of learning problems (Saxe, 1979; Saxe and Shaheen, Note 2) have enabled a different perspective on the character of numeration as a representational system and its relation to a subject's understanding of basic physical and mathematical concepts. In the case of pathogenesis and breakdown, as in the case of development in cross-cultural perspective, the study of numeration is proving to offer a particularly useful, and, to date, minimally exploited focus for the comparative study of cognition.

REFERENCE NOTES

1. Saxe, G.B. The breakdown of numerical cognition following focal brain injury: Two case studies. In preparation.

2. Saxe, G.B., & Shaheen, S. Piagetian theory and the atypical case: An analysis of the developmental Gerstmann Syndrome. Unpublished manuscript, The Children's Hospital Medical Center, Boston, Mass. 1978.

REFERENCES

Piaget, J. Piaget's theory. In P.H. Mussen (Ed.), *Carmichael's manual of child psychology.* New York: Wiley, 1970.

Saxe, G.B. A developmental analysis of notational counting. *Child Development,* 1977, *48,* 1512-1520.
Saxe, G.B. Developmental relations between notational counting and number conservation. *Child Development,* 1979a, in press.
Saxe, G.B. Children's counting: The early formation of numerical symbols. In D. Wolf (Ed.), *New directions in developmental psychology.* (Vol. 1, No. 3). San Francisco: Jossey-Bass, 1979b.
Schaeffer, B., Eggleston, V.H., & Scott, J.L. Number development in young children. *Cognitive Psychology,* 1974, *6,* 357-379.
Vygotsky, L.S. *Thought and language.* Cambridge, Mass.: MIT Press, 1962.
Vygotsky, L.S. *Mind and society.* M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, (Eds.), Cambridge, Mass.: Harvard University Press, 1978.

The Effects of the Cultural Salience of Test Materials on Social Class and Ethnic Differences in Cognitive Performance

Warren Simmons
Laboratory of Comparative Human Cognition

In a series of investigations, Sigel and his associates (Sigel, Anderson, & Shapiro, 1966; Sigel & McBane, 1967; Siegel & Olmstead, 1970) have identified social class differences in the cognitive strategies or styles that are used to classify pictures in a task devised by Sigel. Lower-class children create fewer descriptive and categorical groupings than middle-class children, while making greater use of relational categories in their efforts to classify pictures.

In discussing the origins of this social class difference, Sigel (1970) suggests that differences in categorization style reflect social class differences in representational competence — e.g., the ability to impose structure on two-dimensional material. Furthermore, Sigel suggests that social class differences in representational competence arise from the paucity of distancing experiences available in the homes of children from lower-class backgrounds. Distancing experiences are conceived as activities which heighten differentiation and abstraction (Sigel, 1970).

In a test of Sigel's representational competence hypothesis, I argued that a primary source of social class differences in classificatory competence on the Sigel Conceptual Styles Test (SCST) is the differential salience of the test's items for children from various social class and ethnic backgrounds (cf. Simmons, 1979). I hypothesized that if the cultural salience of pictures on a SCST-like task is equated, there will be minimal social class or ethnic differences in sorting strategies. Culturally salient pictures are defined as those which depict objects, persons, or activities that appear, or are engaged in, frequently and are highly valued within a particular subcultural group.

This notion is similar to the general hypothesis that familiar stimuli will facilitate more sophisticated thinking. For example, Stolz and Tiffany (1972) demonstrated that adults would respond with "child-like" free associations when asked to free associate to low frequency English words. Turgeon and Hill (1977) applied this same line of reasoning to a concept reversal task, modifying developmental trends through changes in the categories that served as stimuli.

The distinction between culturally salient and culturally familiar stimuli, however, is in their relative significance and interest to members of a subcultural group. Stimuli identified as culturally salient are more significant and interesting than ones which are culturally familiar. More will be said about this distinction later.

Since the social class differences found by Sigel map rather well onto differences in categorization patterns used in a variety of developmentally sensitive tasks, I thought it plausible to suppose that Sigel's data might reflect differential knowledge about the stimuli instead of (or in addition to) any general cognitive style associated with early socialization practices.

To test this notion, I presented the SCST and a modified version of this test (referred to as the "Simmons test") to 112 5th- and 6th-grade boys. There were equal numbers of Black lower-class, White lower-class, Black middle-class, and White middle-class subjects in the sample. Social class was determined by information concerning parent's education and occupation according to Reiss' (1961) scale.

Both the SCST and the Simmons test contain 18 items each consisting of 3 pictures. The children were first asked to describe each picture in a particular item, then select 2 pictures that belong together and finally to give their reasons for placing the selected pictures together. If, on a particular item, a child indicated that there were no more reasons for putting a pair of pictures together, I asked if there were any other combinations of 2 pictures for that item that go together and if so, why? Thus, a child could put together as many as 3 different pairs of pictures per item and give any number of reasons per pair. There were no time constraints placed on the children's performance on either test.

The pictures on the SCST and the Simmons test varied in their cultural salience for different population groups. To obtain appropriate stimuli I conducted an interview study in which Black and White, middle- and lower-class children were asked to rate their preference for, and involvement in, academic-cultural activities, games and sports, and white collar, blue collar, and entertainer-athlete occupations. Some of these areas of experience were identified as being culturally salient to White and middle-class subjects, while others were culturally salient for Black and lower-class subjects. Academic-cultural activities and white collar occupations were relatively salient for White and middle-class

0160-3361/79/07-43 $1.00 © LCHC

children, whereas games and sports, blue collar occupations, and entertainer-athlete occupations were culturally salient for Black and lower-class children.

Guided by these findings, the Simmons test was constructed to contain 3 subsets of items (6 items per subset); the pictures for one subset (Sigel-type items) were selected to be culturally salient for White and middle-class children (the pilot data indicated that the pictures on the SCST generally were culturally salient for White and middle-class children). For example, some of the Sigel-type items were pictures of a commercial jetiner, a pilot, and an automobile. A second subset (Simmons-type items) was culturally salient for Black and lower-class children. For example, the items in this subset contained pictures of a professional boxer, a professional baseball player, and a color television set. Finally, a third subset of items (Neutral items) was comprised of stimuli that were not differentially salient for any group of subjects. Sample items in this subset were a picture of an old man playing a fiddle, a smiling young woman, and a young woman playing a guitar.

A within-subjects design was used such that each subject received both the SCST and the Simmons test. The two tests were presented in separate testing sessions that were conducted approximately one week apart. Subjects were individually tested by the author who is a Black male in his middle twenties. Order of presentation of the two tests was counterbalanced across subjects. Subjects' responses were coded by the author and 2 independent raters. Inter-rater agreement ranged from .84 to .87. Since there was an absence of grade-level differences, the data were collapsed across grade-level.

According to Sigel (1970), social class differences in representational competence are responsible for group differences in the use of cognitively more sophisticated sorting strategies. For this reason, the effect of the variations in the cultural salience between the tests, especially as salience was reflected in the use of categorical reasons, is of primary interest.

In comparing the use of categorical reasons (c-reasons) on the SCST and Simmons tests (see Table 1), a significant main effect for Social Class ($F = 4.60$, $df = 1$ & 104, $p < .05$) and a highly significant Ethnicity X Test interaction ($F = 10.10$, $df = 1$ & 104, $p < .002$) were obtained. Referring to Table 1, these effects indicate that across both tests, middle-class children used more categorical reasons than did the lower-class ones. However, whereas on the SCST the White children used more c-reasons than the Black children did, the opposite occurs on the Simmons test.

A further inspection of the means in Table 1 reveals that the current data replicate Sigel's findings vis-à-vis social class differences in the use of c-reasons on the SCST — i.e., relative to their lower-class counterparts, middle-class children tend to use c-reasons more frequently. But the absence of sizeable social class differences in the use of this strategy on the Simmons test is contrary to Sigel's previous evidence. If group differences in basic competence or style are the primary determinant of social class differences in performance on tests of this type, there should not have been any marked test-to-test variance in the pattern of social class differences. The fact that this variance is evident in the present data attests to the extent to which the differences in the cultural salience of the two tests mediates the pattern of social class, as well as ethnic, differences in performance.

Table 1

Mean Use of C-Reasons on Sigel and Simmons Tests

	Sigel Test	
	Lower-Class	Middle-Class
Black	3.185	3.926
White	3.481	5.370
	Simmons Test	
	Lower-Class	Middle-Class
Black	7.07	6.70
White	4.63	6.18

To establish that the varying patterns of performance in the first analysis were a function of the differences in the cultural salience of the 2 tests and not to some other factor(s), comparisons were made of the children's performance on the 3 subsets (Neutral, Sigel-type, & Simmons-type) of items on the Simmons test. Recall that Sigel-type items were considered culturally salient for middle-class and White children, while Simmons-type items were culturally salient for Black and lower-class children. Neutral items were not differentially salient for any group of children.

A 3 (Subsets) × 2 (Social Class) × 2 (Ethnicity) analysis of variance tested the statistical significance of the mean differences in the use of c-reasons on the Simmons test (see Table 2). Though no main effects were apparent, favorable support for the cultural salience hypothesis, however, is generated by the Social Class × Subset ($F = 10.30$, $df = 2$ & 104, $p < .001$) and Ethnicity × Subset interactions ($F = 18.10$, $df = 2$ & 104, $p < .001$) that obtained. In table 2, we see that the former interaction demonstrates that while there are no appreciable social class differences in the use of c-reasons on the Neutral subset, on Sigel-type items middle-class children use more c-reasons than do their lower-class counterparts. In comparison, on Simmons-type items the lower-class children's use of this strategy is greater than that of the middle-class group.

A similar pattern emerges in relation to the Ethnicity × Subset interaction. In this instance, Whites use more c-reasons than do Blacks on Sigel-type items. But relative to the White children, the Black children use more c-reasons on the Simmons-type items. As expected, there were no sizeable ethnic group differences in the use of this strategy on the Neutral items.

There was also a Social Class × Ethnicity ($F = 3.95$, $df = 2$ & 104, $p < .05$) interaction: although White chil-

dren give the greatest number of c-reasons in the middle-class group, in the lower-class group Black children use this strategy more frequently than the White children do.

Table 2
Mean Use of C-Reasons on Simmons Test Subsets

	Lower-Class	Middle-Class
	Sigel-Type	
Black	1.29	2.25
White	2.00	3.29
	Simmons-Type	
Black	4.51	2.85
White	1.96	2.07
	Neutral-Type	
Black	1.15	0.93
White	0.78	0.93

DISCUSSION

Taken together, these classification data show that the cognitive sophistication of the classification strategies used by members of the groups studied varied in accordance with the cultural salience of the pictures they were asked to sort. The fact that the sophistication of the strategies that were used varied with the nature of the material strongly suggests that Sigel et al.'s previous findings may have been dependent upon subject-stimulus relations, rather than a general style located within individuals belonging to certain subcultural groups.

The present data also speak to the merit of stimulus salience in addition to stimulus familiarity as an explanatory concept in classification tasks. Since most of the subjects were able adequately to identify the items — all of which are ubiquitous features of the American scene — there was no basis for arguing that group differences in familiarity were responsible for the results. However, children's ability to provide a ready verbal label for a given stimulus gives little indication of their depth of knowledge regarding the stimulus or the meaning that is extracted from it. In the present study, the subjects were equally familiar with the stimuli, but there were large differences in the extent to which the stimuli were culturally salient for certain subcultural groups.

What is crucial here is the recognition that stimuli that are equally familiar to members of different subcultural groups may vary in their internal organization, and thus meaning, for individuals from different subcultural backgrounds. Because the concept of cultural salience refers to the structure of knowledge, nature of meaning, and interest involved in subject-stimulus relations, it provides a means of linking cultural organization and conceptual behavior that sheer frequency alone does not provide.

The effect of the cultural salience of test materials on social class and ethnic differences in performance also has methodological implications. One of the major caveats in the findings is that characterizations of the ability or response styles of members of various subcultural groups should not be made based upon operations performed on a single set of procedures and materials. Introduction of the Simmons test in the present study, in effect, damaged the validity of previous statements concerning the existence of social class differences in representational competence (cf. Sigel, 1970).

One way to avoid this problem is through the use of within-subjects designs (such as the SCST-Simmons test and Simmons subsets comparisons) that promote attention to interactions, in addition to simple main effects. An additional benefit of this strategy is that it renders neutral such important factors as race and social class of experimenter (cf. Boykin, 1976), and test familiarity (cf. Cole & Means, 1978; Orasanu, 1976). Moreover, the interactions that can occur in a within-subjects design allow for greater separation of the effects of test content on group differences in performance from the effects of group differences in competence.

This research strategy can also be more helpful in distinguishing what has been referred to as the component processes that mediate performance (Cole and Means, 1978). The notion of component processes refers to the strategies and skills that are required to produce the end-product of performance on a particular task. Since previous research has largely been concerned with the relationship between test content and global aspects of performance (Orasanu, 1976), the component processes that determine performance on many tasks remain theoretically and empirically unspecified.

The final methodological point that I'd like to make is that using more than one task to measure a skill broadens the range of application of the processes that are being examined, which in turn enhances the validity of generalizations that are made about how the skill is used on other kinds of tasks. On this issue, however, there still remains a need to study the display of cognitive skills in situations outside of the laboratory and in relation to tasks that are naturally-occurring (cf. Cole and Means, 1978).

As was stated earlier, the present data demonstrate the influence of the content of the materials on subcultural group differences in performance, but in line with the suggestions that have just been made, more needs to be said about the manner in which this influence was exerted. Some nonsystematic observations of how subjects from different subcultural groups approached the task and the experimenter might account for certain aspects of performance that might otherwise be puzzling. Prior to testing I made extensive efforts to establish rapport with the children (and the school staff) by holding informal discussion groups where the children's attitudes toward school regulations, staff, and life in general were broached. The pretext of the testing sessions was a continuation of our previous group meetings on an individual level. Prior to introduc-

ing a test, the session was begun by picking up on a topic from the previous week's conversations. The test was introduced after ten minutes of conversation. In administering the tests to Black and White, lower- and middle-class children, I noticed that there were consistent social class differences in the children's responses to certain features of the testing situation, particularly my probing efforts.

Recall that on each item, subjects were asked to select two pictures that belong together and give any number of reasons for their choice. After a child gave a reason, I would then ask if there were any other reasons for putting the pair in question together. This procedure was followed for each pair in an item that was chosen for all of the children (cf. Simmons, 1979). Middle-class children tended to respond to these probes by confidently assuring me that there were no other reasons, in many cases without even visually attending to the pair again. In contrast, the lower-class children responded to the probes by dutifully searching for (i.e., attending to the pair again) and providing additional reasons. By the expressions on the children's faces and their patterns of attending to the stimuli, it seemed as though the middle-class children interpreted my probe as an effort to trick them into giving a "wrong answer," whereas the lower-class children appeared to view the same probe as feedback about the correctness or incorrectness of their prior response. Support for the latter interpretation can be found in Katz's (1976) research which demonstrates that Black lower-class children tend to be more self-critical in evaluating their own performance than Black middle-class children, though there is some question of the effect of the race of experimenter on this outcome. In this instance, I would argue that the social class differences in response to my probes were due to the different meaning and consequences that tests have for children from varying backgrounds. Sigel (1979) has noted that the structural features of questions (e.g., content, message, and temporal quality) "have to be gauged relative to the child's experiential developmental state" (p. 172). Aside from the child's level of development, one must consider the child's previous experience in similar situations and how this affects his or her interpretation of a question. The different school experience of lower- and middle-class children has been widely documented (cf. Coleman, 1966; Wolf, 1977). Relative to middle-class children, evidence suggests that lower-class children encounter less praise, more criticism, and less informationally-oriented attention (e.g., Rist, 1973) in school environments. The extent to which this circumstance would cause lower-class children to exhibit more doubt and defensive behaviors in testing situations needs to be studied for its implications for social class differences in performance.

The importance of the social aspect of testing situations is seen once again in the Ethnicity × Social Class interaction: in the middle-class group Whites outperform Blacks, while the reverse holds for the lower-class group. Here it seems that the race and social class of the experimenter is again central, particularly as it relates to the race and social class of the subjects. The salient factor here is the social distance between the race and social class of the experimenter and that of the subject, and its effect on performance. Tables 1 and 2 show that the performance of the White lower-class subjects is often less than that of the other groups. One reason for this might have been the White lower-class subject's inability to identify or feel comfortable with the experimenter. I shared in common with the other 3 groups of subjects either race, social class, or both factors. I held neither of these factors in common with the White lower-class subjects. This circumstance might have led them to be more wary of me than the children who were in the other groups, which would account for their almost uniformly (comparatively) low level of response.

In future research, I plan systematically to investigate the issues that have been raised — e.g., the relationship between the social distance between the experimenter and the subject and subject's performance. Among other factors that require study is the relationship between possible subcultural differences in the conception of what constitutes a "smart answer" and group differences in performance. In talks that I've given on this research I've asked members of the audience to sort pictures and provide reasons for their sorts. On several occasions when I asked for evaluation of the reasons there appeared to be ethnic-group differences in what was considered a "good" answer. Until factors such as these are addressed by the literature, source of social class and ethnic differences in classification performance will remain obscure, and global theories suspect.

REFERENCES

Boykin, A.W. Contextual factors and test performance. *Laboratory of Comparative Human Cognition* and *The Institute for Comparative Human Development.* Working Paper No. 2. The Rockefeller University, 1976.

Cole, M., & Means, B. A practicum for comparative cognitive research, Working Paper No. 1. *Laboratory of Comparative Human Cognition.* University of California, San Diego, August 1978.

Coleman, J. S. *Equality of educational opportunity.* Washington, D.C.: U.S. Department of Health, Education, & Welfare, 1966.

Katz, I., Cole, O. J., & Baron, P. M. Self-evaluation, social reinforcement, & academic achievement in Black & White schoolchildren. *Child Development,* 1976, *47,* 368-74.

Orasanu, J. Constraints of text and setting on measurement of mental ability. *Laboratory of Comparative Human Cognition and The Institute for Comparative Human Development, Working Paper No. 3.* New York: The Rockefeller University, 1976.

Reiss, A. J. *Occupation and social status.* New York: Free Press, 1961.

Rist, R. C. *The urban school: A factory for failure.* Cambridge, Mass.: MIT Press, 1973.
Sigel, I. E. The distancing hypothesis: A causal hypothesis for the acquisition of representational thought. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior, 1968: Effects of early experience.* Coral Gables, Florida: University of Miami Press, 1970.
Sigel, I. E. An inquiry into inquiry: Question asking as an instructional model. In L. Katz (Ed.), *Current topics in early childhood education, Vol. 2.* Norwood, N.J.: Ablex, 1979.
Sigel, I., Anderson, L.M., & Shapiro, H. Categorization behavior of low-income and middle-income Negro children. *Journal of Negro Education,* 1966, Jan. 3, 218-29.
Sigel, I. E., & McBane, B. Cognitive competence and level of symbolization among 5 year old children. In J. Hellmuth (Ed.), *The disadvantaged child, Vol. 1.* Seattle: Special Child Publications, 1967.
Sigel, I. E., & Olmstead, P. Modification of classificatory competence and level of representation among lower-class Negro kindergarten children. In A. H. Passow (Ed.), *Reaching the disadvantaged learner.* New York: Teachers College Press, 1970.
Simmons, W. The effects of the cultural salience of test materials on social class and ethnic differences in cognitive performance. Unpublished dissertation. Cornell University, 1979.
Stolz, W. & Tiffany, J. The production of "child-like" associations by adults to unfamiliar adjectives. *Journal of Verbal Learning and Verbal Behavior,* 1972, *2*(1), 38-46.
Turgeon, V. F., & Hill, S. D. A developmental analysis of the formation and use of conceptual categories. *Journal of Experimental Child Psychology,* 1977, *23,* 108-116.
Wolf, R. *Achievement in America.* New York: Teachers College Press, 1977.

A New Multi-Level Analytic Framework for Conducting Cross-Cultural and Socio-Cultural Psychological Research

Leslie J. Fyans, Jr.
Department of Planning and Research
Illinois Office of Education

One primary concern for the cross-cultural study of psychological processes focuses upon the validity of theories and findings drawn from one setting, population, and culture for any other setting, population, and culture (Triandis, 1972). This concern addresses the question of whether statements concerning human behavior should be quite general and at the level of all cultures, or, quite relative and at the level of a specific culture (Bendix, 1963). To insure the veridicality of any cross-cultural statements, the psychologist must be able to investigate the dependent variable(s) of interest at multiple levels (i.e., cultural, subcultural, and personological) so as to be able *to test alternative hypotheses* concerning the cross-cultural level of human behavior.

Unfortunately, traditional cross-cultural research methods have been restricted to investigating a psychological process either at the level of a specific culture (Cole and Scribner, 1974) *or* at the level of all cultures (Rohner, 1975; McClelland, 1961). Eckensbarger (in Nesslereade and Reese, 1973), Berry (1975), and others (Przewerski and Teune, 1970) have called for multiple level research for cross-cultural psychology, *but* have not provided any integrated model from which to carry out a multi-level investigative process.

However, an integrative multiple level model for cross-cultural psychology can be developed (Fyans, Notes 1 and 2) using the logic of generalizability (Cronbach et al., 1972). Central to this model is the capability for the researcher to collect and analyze data from cultural, subsystemic, and personological levels and unconfound (simultaneously) the components of dependent variable(s) variance attributable to culture-specific (nested) and culture-general (crossed) factors. However, there are many other utilities from adopting this Bayesian-Generalizability (Fyans, Note 1) research model. A few of these will be addressed at the end of this article.

This article will discuss this Generalizability approach (Fyans, Note 2) in outlining a sequential decision model which can be used to determine the level of generalizability (or specificity) of human behavior. It is hoped that these presentations will facilitate multiple level investigations by cross-cultural and socio-cultural research.

Sequential Decision Strategy

The sequential decision strategy can be illustrated with a hypothetical cross-cultural study on attributional behavior. A social psychologist could be interested in the attributional behavior of students in situations where they have all experienced success on a task. The social psychologist could be very interested in the cross-cultural generality or specificity of this attributional behavior. Having designed emically equivalent measurements of attribution (etc.) for several cultures, the researcher could be concerned as to whether "success attributions" could be accounted for solely in terms of *personological* differences in achievement motivation level, or if they were dependent upon an individual's *ethnic* membership[1] or *cultural* differences. Likewise, the investigator could be interested in whether there were any differences in "success-attribution" behavior which were associated with *age-cohort* differences. In short, the social psychologist would be investigating how far statements concerning success-attribution behavior could be generalized and the limitations improved by cultural, ethnic, developmental, or personological dif-

0160-3361/79/07-47 $1.00 © LCHC

ferences. Thus, this hypothetical study is a *multiple level* cross-cultural study (Fyans, Note 2) investigating variables at the levels of *individual differences* (achievement motivation), *subcultural* crossed (age cohort), *subcultural nested* (ethnic group), and *systemic* (culture).

The researcher now would engage in the sequential decision strategy with which to decide the veridical level of generality or specificity of attributional behavior. The assessment of the veridical level of generality (or specificity) proceeds through four[2] sequential steps (from personological to systemic), at each step examining the size of the generalization coefficient associated with each variable used in the study. The generalizability coefficients represent the viability of each alternative hypothesis with respect to all of the multiple levels included in the study. The generalizability coefficient for a particular independent variable[3] is simply a ratio of a numerator containing the variable component for the independent variable of interest and a denominator consisting of the same variable component as well as the variance components of any interaction terms that independent variable has with another variable.

Before illustrating the sequential decision strategy relative to our hypothetical study, it is worthwhile to briefly discuss the meaning of generalizability coefficients of various sizes.

Obtaining large generalizability coefficients for a particular independent variable represents high *within-level* correlations among the dependent variable scores for that independent variable. (Fyans and Maehr, in press). Thus, a high generalizability coefficient for the independent variable of "age" would mean that, *within each level* (e.g., elementary, high school, college), there was a high degree of similarity of dependent variable scores, but that there were *large differences* in dependent variable scores *across age levels.* Thus, all discussions or statements concerning the dependent variable would have to be made *specific* to an individual's age level. This high generalizability would thus indicate that there was a high degree of generalizability of dependent variable scores across all other non-age factors included in the study (e.g., achievement levels, ethnicity, cultural membership).

Similarly, a high generalizability coefficient for "culture" would indicate a high *within-culture* correlation among dependent variable scores and it would imply a high degree of generalizability of dependent variable scores across the other non-culture facets included in the study (e.g., achievement motivation, age, ethnicity). Thus, a high generalizability coefficient for "culture" would imply that all statements concerning the dependent variable would have to be made specific to each individual culture.

However, a low generalizability coefficient for a particular independent variable would indicate a lack of specificity in dependent variable scores for that particular independent variable of interest. A low generalizability coefficient could mean that any within-level dependent variable relationship was attenuated by large interaction terms in the denominator of the ratio forming the generalizability coefficient.

Having computed the generalizability coefficients, the researcher would then sequentially determine[2] which level accounted for the most variance of the dependent variable. The decision strategy proceeds via the following four sequential steps.

Step I: The Level of Individual Differences

The researcher would begin at what Przeworski and Teune (1970) define as the level of the *individual actor.* During step I all subjects in a cross-cultural study are treated as if they were from one, single, homogeneous population (Rohner, 1975; Malewski, 1961).

With regard to our hypothetical "success-attribution" study, the researcher would investigate the generalizability coefficient for the independent variable of achievement motivation level. To the extent that a large generalizability coefficient is associated with this individual difference variable (>.80) the researcher could state that success-attributions behavior was accounted for primarily by individual differences in achievement motivation level. Thus, statements concerning success-attribution behavior would be specific to each achievement motivation level but generalizable across age cohorts, ethnic groups, and cultures. This generalizability coefficient would indicate high correlations among the success-attribution scores for any one particular achievement motivation level regardless of their ethnicity or cultural membership. However, it is possible that the generalizability coefficient for the individual differences variable would be low. The researcher would then move to step II.

Step II: The Level of Subcultural Crossed Variables

According to Przeworski and Teune (1970), this level would be most representative of the "most different systems" design. Here the researcher is concerned with generality across cultures and persons especially in exploring dependent variable scores in terms of subcultural cohorts. For our hypothetical study, during step II, the researcher would see if this was associated with the variable of age. This would mean that these statements concerning "success-attribution" behavior would have to be specific to each age level, but that within each age level there would be high average correlations between "success-attribution" scores of individuals regardless of their cultural or ethnic membership or achievement (personological) "type." However, if the generalizability coefficient for age was minimal, the researcher would then proceed to the next steps of the decision sequence.

Step III: The Level of Subcultural Nested Variables

In the previous levels, the investigation was whether statements concerning the dependent variable were generalizable across cultures. Thus, the previous levels were the first steps toward the development of general, culture free laws (Triandis, 1972). However, to the extent that the generalizability coefficients for per-

sonological subcultural-general varibles were not large, the researcher would now have to explicitly recognize some degree of cultural specificity in the explanation of the dependent variable. There are two possible levels of specificity which could alternatively account for the dependent variable.

The first hypothesis of specificity is the more parsimonious one and holds that the dependent variable can be best accounted for in terms of factors whose levels are nested within each specific culture (e.g., ethnic). The alternative hypothesis assumes that within each culture there are so many distinct variables and unique patterns of interaction as to make it impossible to isolate specific nested variables to account for the dependent variable. Thus, this latter hypothesis explicity brings the proper names of each system into the explanation itself (Przeworski and Teune, 1970). It is the purpose of step III to determine the strength of the former, more parsimonious, hypothesis.

In terms of the hypothetical study, the viability of the first hypothesis is tested by the size of the generalizability coefficient for the variable of ethnic groups which are nested within each culture. To the extent that this coefficient is large, the explanation of "success-attribution" behavior would be specific to each ethnic group included in the study. Within each ethnic group there would be a high average correlation of "success-attribution" scores and this generalizability across the achievement motivation and age levels in the study. However, if the generalizability coefficient for ethnic groups was small, the researcher would be forced to move to the final level, step IV, in the decision process.

Step IV: The Level of the Culture Itself

At this level, rather than describing a dependent variable in terms of a few nested variables (step III), the whole of each culture is the focus of analysis. This step assumes that the complexity of cultures is such that it is impossible to define which particular systemic variables direct the cultural influence.

To the extent that the generalizability coefficient for cultures is large, all statements of "success-attributions" must be made specific to each culture in the study, eventuating in the incorporation of the *proper name* of each culture into the explanation of the dependent variable (Przeworski and Teune, 1970). However, with the finding of a large generalizability coefficient for culture in step IV the researcher must further attempt to determine which systemic variables compose the cultural influences on "success attributions." Thus further analysis would be conducted *within* each culture and take the form of an ethnography (Narroll and Cohen, 1970) or normative study (Triandis, 1972). Triandis (1975, 1972) gives the systemic variables of norms, roles, customs, values, attitudes, expectations, ideals, stereotypes, and typical tasks which could be inspected within each culture to determine what contributed to the generalizability coefficient for culture.

Lack of Strong Main Effect Generalizability Coefficients

It is possible that even after discounting the hypothesis associated with steps I, II, and III, the generalizability coefficient for culture itself is not large. There could be two primary reasons for the occurrence of a small (culture) generalizability coefficient during step IV. These two causes consist of, (a) the occurrence of a *strong interaction* effect, and (b) the possibility of *specification error* in the model employed by a cross-cultural researcher.

(a) The Possibility of Strong Interaction Effects

So far in this article, the effects of interaction terms have been relegated to the formulation of the denominator error terms for the generalizability coefficient at each level in the analysis. However, it is possible that the interaction terms themselves (e.g., culture × age or age × achievement motivation) account for much more of the dependent variable variance than do the main effects. This occurrence, in itself, is not troublesome for the cross-cultural researcher, since the researcher uses this information in a manner similar to that of the sequential strategy discussed above. The researcher merely makes any statements concerning the dependent variable specific to the types of factors making up the strong interaction.

Thus, if the interaction effect was composed of individual differences/personological variables (e.g., achievement, etc.), generalizations are still possible across subcultural variables and cultures. Likewise, if the interactions involved subcultural variables (e.g., sex, societal status, town vs. rural residence) generalizations could be made across cultures and individuals classified according to other subcultural variables (e.g., age level). It would also be possible to generalize statements across cultures for an interaction effect made up of a combination of individual differences/personological variables (e.g., achievement, etc.) generalizations are still possible across subcultural variables and cultures. However, if the strong interaction effect was composed in any way of either subcultural-nested variables or of the culture term itself (e.g., ethnic group × age or culture × sex), the statements would have to be made specific to each culture study. Thus, the presence of large generalizability coefficients for interaction terms does not necessarily complicate the analysis since it follows a sequential strategy similar to that for the main effects in the interpretation of generality or specificity.

(b) Specification Error in the Model under Investigation

The problem of specification error occurs when a multiple level model is "underspecified." In other words, the researcher either included the wrong variables or omitted more meaningful variables at each level in the analysis. The omitted variables could be correlated with the independent variable and the residual on the dependent variable. Thus the variables included in the multiple level study did not appreciably account for dependent variable variance. Generally speaking, this

will occur when the researcher has not spent much time prior to data collection in developing the theory to be tested and all of its ramifications. Thus with more careful thought and study, the researcher would have included more theoretically meaningful variables to represent the levels in the multiple-level analysis. Having included these variables in the study will have resulted in the veridical level of generality or specificity of human behavior being identified by the model during the sequential decision process.

CONCLUSION

This paper has presented several facets of a new paradigm from which to conduct cross-cultural psychological research. Unlike the past models guiding cross-cultural psychological research, the new paradigm facilitates multiple level investigations by incorporating both culture-specific (nested) and culture-general (crossed) independent variables within its partially-hierarchical framework. Based upon the generalizability analysis, this model generates sequential tests of alternative hypotheses so that the veridical level of generality or specificity of human behavior can be succinctly assessed. Having empirically determined the veridical level of generality, the researcher can then employ a regression equation which is congruent to the level of generality found. This regression equation integrates all within culture, between culture, and individual differences information into one prediction equation. This regressed universe score approach is more reliable (Shigamasu, 1976) and less susceptible to the alternative hypotheses which plague cross-cultural research (Brislin, Lonner, and Thorndike, 1973; Triandis, 1972). Similarly, the Bayesian approach incorporates information concerning the level of generalizability in tests between any or all of the cultural means on the dependent variable. The new approach is open to any number of crossed and nested variables. Indeed the author is currently collaborating with Maehr and Silili (Note 3) in employing this paradigm to analyze cross-cultural data on concepts of achievement, success, and motivation which entails two nested and seven crossed factors, and with Sprague, Maehr, and Cohen (Note 4) on a cross-cultural study of hyperactivity in children. It is hoped that this new approach will be adopted by other cross-cultural researchers in order that multiple-level investigations can be accomplished and the veridical nature of human behavior properly assessed.

REFERENCE NOTES

1. Fyans, L.J., Jr. A new paradigm for cross-cultural psychological research. Unpublished Doctoral Dissertation. Doctoral Committee members are Maurice Tatsuoka, Robert Linn, Martin Maehr, Harry Triandis, Helen Farmer, and Charles Lewis.

2. Fyans, L.J., Jr. A new paradigm for cross-cultural psychological research. Paper presented at the Psychometric Society, August 1978, McMaster University, Ontario, Canada, and at the American Psychological Association Convention, August 1977, San Francisco.

3. Fyans, L.J., Jr., Maehr, M.L., & Silili, F. A cross-cultural study of achievement related concepts: An analysis of thirty cultures. Unpublished manuscript, Institute for Child Behavior and Development. University of Illinois, Urbana-Champaign.

4. Sprague, R., Maehr, M.L., Fyans, L.J., Jr. & Cohen, M. The hyperactive child from a cross-cultural perspective. Study in progress, Institute for Child Behavior and Development, University of Illinois, Urbana-Champaign.

REFERENCES

Bendix, R. Concepts and generalization in comparative sociological studies. *American Sociological Review,* 1963, *28,* 532.

Berry, J.W. Ecology, cultural adaptation, and psychological differentiation: Traditional patterning and acculturative stress. in R.W. Brislin, S. Bochner, & W. J. Lonner (Eds.), *Cross-cultural perspectives on learning.* New York: Wiley, 1975.

Bowers, K.S. Situationism in psychology: An analysis and a critique. *Psychological Review,* 1973, *80,* 307-336.

Brislin, R.W., Lonner, W.J., & Thorndike, R.M. *Cross-cultural research methods.* New York: Wiley, 1973.

Campbell, D.T. Natural selection as an epistemological model. In R. Naroll & R. Cohen (Eds.), *A handbook of method in cultural anthropology.* New York: Columbia University Press, 1973.

Cohen, R. Generalizations in ethnology. In R. Naroll, & R. Cohen (Eds.), *A handbook of method in cultural anthropology.* New York: Columbia University Press, 1973

Cole, M., & Scribner, S. *Culture and thought: A psychological introduction.* New York: Wiley, 1974.

Cronbach, L.J., Gleser, G.C., Nanda, H., & Rajaratnam, N. The dependability of behavioral measurements. New York: Wiley, 1972.

Eckensbarger, L.H. Methodological issues of cross-cultural research in developmental psychology. In J.R. Nesslereade & H.W. Reese (Eds.), *Lifespan developmental psychology: Methodological issues.* New York: Academic Press, 1973.

Haller, A.O., Lewis, D.I., & Ishine, I. The hypothesis of intersocial similarity in occupational prestige. *American Journal of Sociology,* 1966, *71.*

Fyans, L.J., Jr., & Maehr, M.L. Attributional task selection and achievement. *Journal of Educational Psychology,* in press.

Malewski, A. Two models of sociology. *Studia Socjologiczne,* 1961, *1.*

Miller, D.R. The personality as a system. In R. Naroll & R. Cohen (Eds.), *A handbook of method in cultural anthropology.* New York: Columbia University Press, 1973.

McClelland, D.C. *The achieving society.* New York: Free Press, 1961.
Murdock, G.T., et al. *Outline of cultural materials.* New Haven, Conn.: HRAF Press, 1967.
Naroll, R., & Cohen, R. (Eds.), *A handbook on method in cultural anthropology.* New York: Columbia University Press, 1973.
Przeworski, A., (Teune, H. *The logic of comparative social inquiry.* New York: Wiley, 1970.
Rohner, R.P. Parental acceptance-rejection and personality development: A universalist approach to behavioral science. In R.W. Brislin, S. Bochner, & W. J. Conner (Eds.), *Cross-cultural perspectives on learning.* New York: Wiley, 1975.
Rosch, E. Universals and culture specifics in human categorization. In R.W. Brislin, S. Bochner, & W.J. Conner (Eds.), *Cross-cultural perspectives on learning.* New York: Wiley, 1975.
Shigamasu, K. Development and validation of a simplified m-group regression model. *Journal of Educational Statistics,* 1976, *1,* 157-180.
Singer, M. A survey of culture and personality theory and research. In B. Kaplan (Ed.), *Studying personality cross-culturally.* Evanston, Ill.: Row, Peterson, 1961.
Tanter, R. Toward a theory of political development. In R. Naroll & R. Cohen (Eds.), *A handbook on method in cultural anthropology.* New York: Columbia University Press, 1973.
Triandis, H.C. *The analysis of subjective culture.* New York: Wiley, 1972.
Triandis, H.C. Culture Training, Cognitive Complexity and Interpersonal Attitudes. In R.W. Brislin, S. Bochner, & W.J. Lonner (Eds.), *Cross-cultural perspectives on learning.* New York: Wiley, 1975.

FOOTNOTES

[1]Assuming for the moment several ethnic groups within each culture and different culture to culture.

[2]Actually, the generalizability coefficients from all multiple levels will be computed simultaneously, but can be treated successively in a sequential decision process.

[3]Generalizability coefficients can also be calculated for each interaction term, as will be discussed shortly.

ANNOTATED BIBLIOGRAPHY

Kintsch, W., & Greene, E. The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes,* 1978, *1,* 1-13

Kintsch and Greene report on experiments designed to show the effect of story schemata, which are culture-specific, on the comprehension and recall of texts. American students were given texts with a familiar, European-based structure (native story schema), and an unfamiliar, non-European structure (foreign story schema). While students did not rate individual sentences from the foreign passage as significantly different from those of the native passage in comprehensibility, imagery, or bizarreness, when they read and rated the entire passages, the foreign stories were considered much more bizarre than the native stories. Similarly, when subjects were asked to write summaries of the texts, their summaries of the foreign stories were ranked significantly lower by independent judges. Using a sequential recall methodology, the authors found that even when accuracy was stressed, there was a significant difference in how much of the gist of the native and foreign stories was recalled after five retellings. They suggest that story comprehension and recall are functions of familiarity with the structural schemata.

Steffensen, Margaret S., Jogdeo, Chitra, & Anderson, Richard C. A cross-cultural perspective on reading comprehension. Technical Report #97, Center for the Study of Reading, University of Illinois at Urbana-Champaign.

In another cross-cultural study, Steffensen, Jogdeo, and Anderson report on an experiment designed to show the effect of content schemata, which are also culture-specific, on the comprehension and recall of texts. American and (Asian) Indian subjects were given two texts that were culturally loaded: one with American-based content, and one with Indian-based content. Subjects read their native passage more rapidly, and recalled more information from it. They produced many culturally appropriate elaborations when they recalled their native passage, and also produced many culturally based distortions when they recalled the foreign passage.

These studies need replication with subjects from dominant cultures, and from cultures embedded and encapsulated such as inner-city blacks, American Indians, and Chinese from Chinatowns. It can still be said that the pervasive influence that the information assumed and the structure used by the writer appear to have on both comprehension and memory has implications for the selection of reading materials for both foreign and minority students. The culture structure and content of the materials should match as closely as possible the background of the students. It is only when students have some proficiency in reading that materials based on the target culture should be introduced.

Chitra Jogdeo

MOVING?

Please give us as much advance notice as possible and avoid missing an issue of the Newsletter.

Subscription Form

Name _____

Address _____

_____ Zip _____

Please enter my subscription to *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition.*

I am enclosing $ for years at $5.00 per year.

Please make your checks payable to LCHC NEWSLETTER and mail them to:

Laboratory of Comparative Human Cognition, D-003
University of California, San Diego
La Jolla, CA 92093