FINAL REPORT

Project No. 3-2681

Grant No. NE-G-00-3-0101

The Pupillary Response as an Index of Cognitive Processing in Mentally Retarded Persons

Joseph A. Buckhalt[1]

George Peabody College

Nashville, Tennessee 37203

January 16, 1975

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

U.S. DEPARTMENT OF

HEALTH, EDUCATION, AND WELFARE

NATIONAL INSTITUTE OF EDUCATION

OFFICE OF RESEARCH GRANTS

[1]Present Address: School of Education, University of Southern California, Los Angeles, California, 90007.

ED123863

FINAL REPORT

Project No. 3-2681

Grant No. NE-G-00-3-0101

The Pupillary Response as an Index of Cognitive Processing in Mentally Retarded Persons

Joseph A. Buckhalt[1]

George Peabody College

Nashville, Tennessee 37203

January 16, 1975

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

U.S. DEPARTMENT OF

HEALTH, EDUCATION, AND WELFARE

NATIONAL INSTITUTE OF EDUCATION

OFFICE OF RESEARCH GRANTS

[1]Present Address: School of Education, University of Southern California, Los Angeles, California, 90007.

EC090233

INTRODUCTION

I. An active processing strategies model of memory.

The construct of intelligence has for a long time rested to a great extent not upon learning ability or acquisition *per se*, but upon memory, or the retention of information presumably learned. With no knowledge of whether or not information was ever presented to a subject in the first place, and no knowledge of the original level of learning if it was presented, we have, on our best measures of intelligence, proceeded to make assumptions about intelligence from measures of retention. Consider, for example, the Information, Similarities, and Vocabulary subtests of the Wechsler intelligence scales. Those three subtests obviously rely upon the testee's memory for information to which he is assumed to have been exposed. A little reflection reveals, however, that practically all other specific requirements of tests of intelligence rest upon similar assumptions and depend upon retention of verbal information or perceptual-motor skills.

As Hunt (1961) has pointed out, the idea of fixed intelligence has been one of the most ubiquitous concepts in the field of psychology. Most professionals in psychology and education believed for a long time, and many still believe, that intellectual abilities are more or less "given" at birth, and fairly fixed or relatively impervious to any attempts to modify them in any significant way. It is no wonder, then, considering that memory has played such an integral part in our concept of intelligence, that memory, too, has been thought of as essentially fixed. The ability to remember things has most often been thought of as a "gift", and, like intelligence, considered normally distributed in the population.

Hunt (1961) presented a most convincing argument against the concept of fixed intelligence. If we accept his conclusion as a premise, and consider how large a part memory plays in our concept of intelligence, it becomes reasonable to assume that memory processes are amenable to environmental intervention. An active processing strategies model of memory assumes that the ultimate recall of any information is largely a function of what the organism actively does when he is initially confronted with the stimulus complex containing that information. Memorization is not an automatic process, nor does it always necessarily involve the same types of cognitive activities. The act of committing something to memory, rather, involves one or more of many possible cognitive strategies which are put to use in an intentional manner. Flavell (1970) has been one of the proponents of such a view of voluntary memory in his numerous suggestions that the development of memory with age consists of learning, and learning to use, various acquisition strategies. There is now an abundance of empirical evidence to support such a notion (Belmont & Butterfield, 1969, 1971; Flavell, 1970; Flavell, Friedrichs, & Hoyt, 1970; Hagen, 1971; Neimark, Slotnick, & Ulrich, 1971; Appel, Cooper, McCarrell, Sims-Knight, Yussen, & Flavell, 1972). Appel et al. (1972), for example, have demonstrated that deliberate memorizing only gradually emerges as a separate and distinct form of cognitive encounter with the environment. Memorizing necessarily includes perceptual contact with the environment, but it goes beyond the developmentally earlier act of perception. Neimark et al. (1971) found that when identical materials and instructions are presented to children of different ages, the increased recall on the part of the older children is a function of qualitative changes in what the children do with the materials at the time of initial presentation.'

With respect to mentally retarded persons, or for that matter, any persons who have trouble recalling information, the preceding model suggests several possible reasons for their poor recall. One possibility is that they have never learned memorization strategies. It is interesting that while we have endeavored to teach children content-oriented subjects like history and geography and sill-oriented subjects like reading and arithmetic, skills like strategies for remembering things have been only incidentally considered. Flavell (1971) concludes his analysis of the development of memory with a quote from Adams (1967) which bears repeating here.

> "We have yet to appraise thoroughly the variables for teaching effective mediators, but the practical implications of this method are large. There is no reason why schoolteachers of future generations should not show students ways of learning materials that will result in their high recall. At present, students are given materials for learning and left to their own memory devices. How much better it would be if an instructor told the students about proved mnemonic devices and saw that they used them in systematic ways (Admas, 1967, p. 134)".

Beyond the possibility that some retarded persons have never learned memorization strategies, there is the possibility that they have learned some strategies but not the most efficient and productive ones. Furthermore, there is the issue of knowing when and where to use the strategies which one does have, as well as the issue of propensity to use those strategies.

In addition to the studies already cited as empirical evidence for the validity of an active processing strategies model of memory, some researchers have attempted to specify what these strategies are. When a person is confronted with information to be memorized, there are at least three different processing strategies which have been found to facilitate recall. One can "rehearse" the information, or simply repeat it over and over again so that it remains in immediate consciousness for a longer period of time; one can "organize" the input in some manner, or one can "elaborate" the input--add something to it. The work of Belmont & Butterfield (1971), stemming from the theoretical speculations of Ellis (1970) concerning the function of rehearsal processes in memory of retarded persons, has shown that when retarded persons are taught to apply active rehearsal strategies, their recall is facilitated. Some of the evidence (e.g. Spitz, 1972; Spitz & Webreck, 1972, both of which are extensions of Spitz's (1966) theoretical speculations about the function of input organizational capacity of retarded persons) shows that retarded persons can markedly improve their recall performance if they are presented with stimulus materials which have highly salient organizational cues. With respect to elaboration, several studies (e.g. MacMillan, 1970; Jensen & Rohwer, 1963; Turnure & Walsh, 1971; Taylor, Josberger, & Knowlton, 1972) have found that when retarded subjects are taught to provide mediating links (i.e., to "elaborate") for paired associates to be remembered, their subsequent recall is facilitated.

All of these recent findings, taken together, are extremely encouraging, especially in light of Belmont's (1966) opinion that no conclusive evidence has been found that retention or memory per se is deficient in retarded populations. Belmont and Butterfield (1969) further conclude that the literature contains no compelling evidence that forgetting rate decreases with either age or intelligence.

One task before us, as researchers, is to continue to demonstrate that active processing strategies can be taught which facilitate the recall of information. More importantly, however, we must attend to one of Flavell's (1971) suggestions that the next logical direction future research on the issue must take is toward facilitation of memory on practical, functional, "real-world" types of tasks. We must become aware of the types of information it is important for retarded persons and children to be taught to remember, and we must then attempt to teach processing strategies which can be tailored to the various types of information.

II. Measures of Aquisitional Strategies

A number of different measures of acquisitional processes have been used in developmental studies of memory. Neimark, Slotnick, and Ulrich (1971), for example, found consistent differences in the ways older and younger children organized pictures during study sessions. This measure was found to correlate positively with amount of material recalled, and it was inferred that older subjects' greater level of recall was in part due to their greater propensity to organize materials spontaneously during acquisition. In several studies, verbal rehearsal during acquisition has been observed directly, and developmental differences in this measure have been related to the development of better recall (Flavell, Beach, & Chinsky, 1966; Keeney, Cannizzo, & Flavell, 1967). Belmont and Butterfield (1971a, 1971b) relied upon the analysis of hesitation patterns on serial learning tasks to infer that retarded subjects were not rehearsing in the same manner as nonretarded subjects. Ellis' (1970) conclusions concerning a "rehearsal deficit" in retarded subjects were the result of inferences made from different serial position curves for retarded and nonretarded groups. In some studies, nonverbal acquisitional strategies like pointing have been observed directly (Daehler, Horowitz, Wynns, & Flavell, 1969). In a recent developmental study, Appel, Cooper, McCarrell, Sims-Knight, Yussen, & Flavell (1972) used overt behavioral evidence (naming; lip movement; cumulative naming as well as an inference drawn from differential recall performance by older and younger groups as a function of instructional condition to support Flavell's (1970) hypotheses concerning memory development. For younger subjects, no differences in recall resulted from instructions to remember a series of pictures when compared with instructions merely to look at the pictures. Recall performance was better in the "remember" condition for older subjects, however, and this fact strengthened the inference that active strategies employed during study periods significantly affect recall, and develop with age.

The understanding of developmental aspects of acquisitional processes has, then, rested upon several kinds of empirical evidence differing in degree and type of inference involved. The diversity of the types of measures used to support the contention that acquisitional processes develop with age contributes substantially to the validity of that contention. In one of his methodological suggestions for future research, Flavell (1970) suggests that researchers"...investigate the occurrence-nonoccurrence and the structure of ordinarily covert mediational activities by rendering them observable... (p.207)." Requiring subjects to rehearse overtly during acquisition is an example of one method of doing this, but the analysis of physiological changes during processing is felt to be a technique which offers several advantages beyond measures which have been used previously.

In studies involving adult subjects, one physiological measure, the pupillary response, has been demonstrated to be a useful index of covert mental processes (Hess & Polt, 1964; Kahnemah & Beatty, 1966; Stanners & Headley, 1972). Such a measure is beneficial because, unlike many inferential measures, it can index acquisitional processes as they are occurring. The measure also does not require forcing the subject to perform overtly operations he might ordinarily perform in a covert manner. Also, since the measure has been shown to reflect nonverbal processes such as imagery (Paivio & Simpson, 1966), it is applicable to the measurement of covert mental operations which are perhaps unobservable in any more direct manner.

Hess's (1965) article which summarized his research for several years on the pupillary response was instrumental in focusing a lot of subsequent attention on the measure, although his was not the first research on the significance of the pupillary response to "psychological" stimuli (see Goldwater, 1972 or Lowenfeld, 1958 for reviews which present broad historical perspectives). One aspect of the pupillary response brought out by Hess (1965) was that it appeared to be sensitive to tasks felt to be emotionally arousing. A lot of subsequent research has found the pupillary response to be consistently related to a construct which has been called at different times, "interest", "emotion", "affect", "motivation", or "preference". Nunnally, Knott, Duchnowski, and Parker (1967) looked at the pupillary response to five types of stimulation which were felt to all somehow relate to the overriding construct of "activation" and found support for the idea that the pupillary response is one component of the orienting response and that the pupil will even dilate to a sudden increase in illumination, thus overriding the normal light reflex. In looking at all the research on "non-cognitive" aspects of the pupillary response, it seems undeniable that dilations in certain situations are most accurately described as being part of, or due to, a general activation, or orienting response.

Hess's (1965) article also presented evidence that the pupillary response can be a measure of more than arousal. In one of the studies reported (Hess & Polt, 1964), two basic aspects of the pupillary response during a task involving "cognitive processing" or "mental effort" were detailed. The first aspect is what Kahneman (Kahneman & Beatty, 1966; Kahneman, Onuska, & Wolman, 1968; Kahneman & Peavler, 1969) has come to refer to as a loading-unloading phenomenon. On a digit span task, for example, the pupil gradually dilates ("loads") during digit presentation and subsequently constricts ("unloads") as the subject reports the recalled digits. The second aspect concerns a problem difficulty factor. It was shown in the Hess & Polt (1964) study that multiplication tasks independently rated as more difficult got greater and greater degrees of pupillary dilation. This relationship between difficulty of problem presented and mean extent of dilation has been replicated with a variety of different types of mental tasks. Kahneman and

Beatty (1966) found progressively greater degrees of dilation for memory tasks involving digit strings of different lengths and for digits, words, and digit transformations. Other studies supporting the relationship are Shaefer, Ferguson, Klein, and Rawson (1968); Payne, Parry, and Harasymiw (1968); Bradshaw (1968); Kahneman, Onuska, and Wolman (1968); Paivio and Simpson (1966).

It appears that pupillary changes in the preceding types of tasks are reflective not of respondent, or reflex behavior, but of voluntary, directed behavior which could be considered essentially no different from more directly observable operant behavior. The pupillary response seems to be consistent with theories of what the subject is trying to "do" to the stimuli presented. As Kahneman (Kahneman & Wright, 1971) has pointed out, it is incumbent upon the investigator to have some theory of what the subject is doing when the pupillary response is recorded if he wished to make assumptions about the response as indicative of a specific psychological process. Kahneman's expressed purpose is to build construct validity for the construct of rehearsal by bringing together a tentative theory and a tentative measure (the pupillary response) of rehearsal. The idea is, of course, that once the validity of the measure as an index of processing load or mental effort has been established, it can be used with fewer restrictions and qualifications to answer questions about the process it supposedly mirrors.

A central methodological issue has been whether or not the pupillary measures of "cognitive processing" on one hand, and "arousal" on the other are confounded in any given experimental task. Several studies (Kahneman & Beatty, 1967; Kahneman & Peavler, 1969; Johnson, 1971; Kahneman & Wright, 1971; Stanners & Headley, 1972) have attempted to unravel this complex issue. Every study but one (Kahneman & Peavler, 1969), however, has involved tasks differing in difficulty. As an example, the Kahneman and Beatty (1967) study was concerned with pupillary responses in a pitch discrimination task. It was found that as the comparison tone got more and more similar to the standard tone (i.e., the task got more difficult), dilations were correspondingly greater. Kahneman interpreted the dilations as indicating increased processing load during the more difficult discriminations, saying, "the magnitude of the responses that have been described here is clearly not determined by the arousing characteristics of any stimulus, rather it corresponds to what the organism does with the information conveyed by a particular stimulus". (p. 104) The alternative explanation is, however, that a more difficult discrimination could be accompanied by increased arousal and that the arousal construct can account for the data just as adequately. Kahneman recognized this argument, but feels that, "...proponents of such a view may have to broaden the concept of anxiety so much as to make it either virtually meaningless or else a synonym of processing load". (p. 104).

The Kahneman and Peavler (1969) study appears to be the strongest support for the "processing load" hypothesis, since that study involved tasks of equal difficulty. The idea was to see if differential incentive values placed upon pairs to be associated would evidence different pupillary responses which would correlate highly with items actually learned. In the paradigm, during the study trials the stimulus cued the subject as to whether the pair to be learned was a high incentive (i.e., learning it would pay off more) or a low incentive pair. From the "arousal" point of view, subjects should have been differentially aroused from the onset of the stimulus. High incentive stimuli should have evoked greater pupillary dilations due to greater arousal. From the "processing" point of view, the differential dilation should not have come until both stimulus pairs had been exposed, since the dilation should be due to effort exerted while trying to form the necessary association. The results favored the "processing" interpretation.

If use of pupillary measures on cognitive tasks could only result in more information concerning problem difficulty, its continued use would be difficult to defend, in light of the availability of simpler, equally informative measures such as error rates or response latencies. In several studies, however, the measurement of pupillary activity has provided information about covert mental operations which is inaccessible by any other means (Kahneman & Wright, 1971; Stanners & Headley, 1972; Stanners, Headley & Clark, 1972; Wright & Kahneman, 1972). Kahneman and Wright (1971), for example, were able to confirm several hypotheses concerning the nature of covert rehearsal as a function of varying task demands (whole vs. probed recall; long vs. short retention intervals) using pupillary measures. In that study, as in many others, the alternative means of elucidating the nature of rehearsal activity was to ask subjects for introspective reports. Subjects in such experiments typically have no clear idea of the effects which subtle task demand changes have upon their processing strategies.

Few studies have been conducted which have used pupillary activity as an index of covert processes during cognitive tasks with children or retarded persons as subjects. In one of the only studies which has done so, the effects of arithmetic difficulty on pupillary dilations in normal and retarded children appeared to support the feasibility of using the pupillary response as an empirical means of studying individual differences in cognitive processing (Boersma, Wilton, Barham, & Muir, 1970). The present studies represent a series of propaedeutic studies to confirm the sensitivity of the measure to individual differences in processing ability.

GENERAL PROCEDURES

Subjects

All subjects had relatively light-colored irises (blue,green, or light brown). This restriction was necessary because of limitations in the data-gathering technique. Maximal contrast between iris color and pupil color (black) is necessary to provide photographs clear enough to obtain pupillary measurements.

Apparatus Materials

As subjects listened to stimulus presentations, they rested their chins on a padded, adjustable chin rest, and gazed into a viewing box measuring 58cm X 60 cm X 68 cm. A Bolex H16 16mm movie camera equipped with a 100mm macro-lens and 15mm of extension tubes is mounted outside the viewing box on the left side with the lens protruding into the box. Mounted inside the box is a small (10 cm X 16 cm) adjustable optical quality mirror which reflects the image of the subject's right eye into the camera lens. The camera is driven by a 15 rpm electric motor at a film speed of two frames per second. A frame counter geared to operate directly from the motor drive insures precise coordination of film frames with stimulus presentation.

A translucent screen covering one side of the viewing box is illuminated with two 15 watt fluorescent bulbs such that the level of illumination at the subject's eye is approximately 65 ftc. This level of illumination is sufficient to obtain clear photographs with Kodak Tri-X Reversal film at an f-setting of 2.8.

Following data collection, film was removed from the camera and developed as negatives. A frame-by-frame measurement of pupil size was accomplished in the following manner: Using a microfiche reader, each frame was magnified approximately 15 times actual size and the resulting vertical diameter of the magnified pupil was measured with a millimeter ruler. Only those frames in which the subject was looking directly ahead towards the center of the screen were measured. Shifting of focus to nearer points or darker areas within the viewing box results in artifactual dilation of the pupil.

Average pupil size during a 10-second period before presentation of the first stimulus constituted a baseline measurement. Pupil sizes during stimulus presentation were converted to deviations (positive or negative) from each subject's individual baseline measure. In order to compress the pupillary data, deviations from baseline were averaged every five frames, with no average being made on fewer then three good measurements within the block of five frames (BFs). Pupillary deviation scores were expressed in units of tenths of millimeters of magnified pupil size. For example, if a subject's baseline average was 20.0 mm, and his pupil size for the first block of frames was 21.5 mm, his first pupillary deviation score would be +15. If his pupil size for the next block of frames was 18.5 mm, his second pupillary deviation score would be -15.

Experiment I: A Developmental Study of Pupil Size During the Presentation of Words for Free Recall

In the present study, analysis of pupillary changes during the presentation of stimulus materials were related to age and instructional condition using a task similar to that used by Appel, Cooper, McCarrell, Sims-Knight, Yussen, & Flavell (1972). If the pupillary evidence converges with the overt behavioral and inferential evidence reported in that study, older children should exhibit greater pupillary dilation in a condition in which they are instructed to remember than when they are instructed only to listen to the stimuli. Younger children and retarded children, on the other hand, should exhibit less differences in dilation as a function of instructional condition. This hypothesis is derived from theoretical considerations regarding both memory development and the nature of pupillary changes during cognitive tasks.

Method

Subjects

Thirty children enrolled in public school classes served as subjects for this study. The three experimental groups were as follows: ten children (5 males and 5 females) from first-grade classrooms; ten children (6 males and 4 females) from fifth-grade classrooms; ten children (4 males and 6 females) from fifth-grade classes for the educable mentally retarded. Means and standard deviations of CA and IQ data are presented in Table 1. IQ scores of the retarded children were taken from

---

Insert Table 1
here

---

school records, and were all based on administrations of either the WISC or the Stanford-Binet. No score was based on a test administration which was more than two years old.

Procedure

Each subject was taken individually by the experimenter to the room where the experiment was conducted. After both subject and experimenter were seated, the experimenter began explanation of the task.

Two tape-recorded lists of nine words were presented to each subject through headphones at a rate of one word every five seconds. Both lists consisted of familiar nouns, three words from each of three categories. List I words were hand, foot, ear, apple, orange, peach, plate, spoon, and fork. List II words were bicycle, truck, bus, cat, pig, cow, belt, shirt, and coat. The presentation order of words for each list was pseudorandom and fixed, with the restriction that no two words from the same category were presented successively.

All subjects listened to two lists of words: one list under a "remember" instruction,and one list under a "listen" instruction. The order of instructional conditions and specific lists (I or II) were counter-balanced over all subjects. The two instructional conditions were designed to be similar to those used by Appel, Cooper, McCarrell, Sims-Knight, Yussen, and Flavell (1972). One obvious difference was that inasmuch as the present study involved an auditory rather than a visual task, the instructions were "listen" and "remember" rather than "look" and "remember". Instructions varied slightly depending upon whether the subject was in a remember-listen or a listen-remember group. Subjects in the listen-remember group received the following instructions:

> Listen: "I have some words I want you to listen to. I want to know if you can hear all these words clearly. After you've heard all of them, I'll ask you if you could hear them all clearly."
>
> Remember: "I have some more words I want you to listen to. This time, though, I want you to try to remember the words you hear. After you've heard them all, I'll ask you to tell me what words you heard."

Subjects in the remember-listen group received the following instructions:

> Remember: "I have some words I want you to listen to. I want to know how many of them you can remember. After you've heard them all, I'll ask you to tell me what words you heard."
>
> Listen: "I have some more words I want you to listen to. This time, though, I just want to know if you can hear all of the words clearly. I'm going to use these words with some children, and I want you to help me check them. After you've heard them all, I'll ask you if you could hear them all clearly."

All subjects were pretested for their understanding of each of the two tasks. For the listen task, the subject was given an instruction to listen, and was given two short trials of three words each. On trial one, the experimenter presented two of the words clearly, and one word was spoken in a barely audible whisper. On trial two, all three words were spoken clearly. In order for a subject to pass this pretest, he was required to answer correctly the question, "Could you hear all the words clearly?" for both trials. For the remember task, the subject was given an instruction to remember, and was given one trial of three words. The requirement for passing this pretest was the correct repetition of all three words when asked to recall them. No subjects were excluded from the experiment for failure to pass the pretests.

Upon completion of the pretests, each subject's head was aligned precisely on the chin rest such that the right eye was within the camera's field, and the proper experimental instruction was given. In addition, subjects were instructed to hold their heads as still as possible and to look at the cross in the center of the screen as they listened to the words.

The camera motor and the tape recorder were switched on simultaneously, and both were switched off at the completion of each list. Following presentation of each list, recall for the words was tested. For the listen condition, in which no instruction to remember was given, the experimenter said, "I know, I didn't ask you to try to remember what the words were, but just for fun let's see if you can remember any of them."

Because eventual analyses of pupillary data were based on natural logarithm ($\log_e$) transformations, a constant was added to all pupil size deviation scores to convert them to positive numbers. Since word presentation for each list lasted 45 seconds (nine words; one word every 5 seconds), and film speed was two frames per second, for each list a subject's pupil was photographed 90 times during word presentation. Averaging every five frames, this figure was reduced to 18 pupillary deviation scores per subject. Two other indices of pupillary activity were used. Since all subjects showed initial dilation (positive pupillary deviation scores) during word presentation, the magnitude of each subject's greatest initial dilation, and the time to that first dilation "peak" were measured.

Results

In order to achieve the desired sample of 30 subjects, 10 in each group, a total of 47 subjects were tested. Of these, pupillary scores from 17 subjects could not be used for various reasons. In some cases, too many film frames were unscoreable because subjects did not follow instructions to gaze at the center of the screen during word presentation. In other cases, excessive head movement caused blurring of film, or the head was moved forward or backward out of the camera's range of focus. The criterion set for discarding subjects' data was this: if 30% or more of the film frames were judged unscoreable by either of the scorers, that subject's pupillary measures were not used. Of the 17 subjects whose pupillary measures did not meet that criterion, nine were from the first grade group, three were from the fifth grade group, and five were from the retarded group. Mean recall scores for the whole (N=47) and the restricted (N=30) samples are presented in Table 2.

Insert Table 2
about here

To determine the effect of excluding subjects on the recall scores, mean scores were compared with *t*-tests by group and condition. Since no significant differences were found, only the recall scores of the 30 subjects in the restricted sample were included in subsequent analyses.

Recall Data

A 3 (Groups) X 2 (Instructional Conditions) X 2 (Order of

Conditions) analysis of variance was performed on the number of words recalled. While no significant main effects were revealed, the predicted Group X Instructional Condition effect, seems to be due to the fact that while performance of the fifth grade group in the Remember condition was superior to their performance in the Listen condition, performance of the other two groups was comparable in both conditions.

Insert Table 3
about here

Degree of clustering was assessed with the following index[1] which related the amount of clustering exhibited to the amount possible relative to the number of words recalled:

$$\text{Clustering index} = \frac{\text{number of common category words juxtaposed in recall}}{\text{total number of words recalled}} \times 100$$

For example, if five words were recalled, a clustering index of 100 would require all three words from one category and two of the words from an additional category juxtaposed in recall.

The effects of experimental conditions on clustering were determined by means of a 3 (Groups) X 2 (Instructional Conditions) X 2 (Order of Conditions) analysis of variance. The only significant effect was the Group main effect ($F = 5.37$, 2/24 df, $p < .05$). No other main effects or interactions were significant. Inspection of the group means presented in Table 3 shows that clustering was greatest in the first grade group, followed by the retarded group, and the fifth grade group. Since no main effects or interactions involving the factors of Instructional Condition or Order of Conditions were significant, means in

Table 3 were collapsed over those two factors.

The degree of cumulative recall was assessed with the following index which is similar to the clustering index:

$$\text{Cumulative Recall Index} = \frac{\text{number of successively presented words juxtaposed in recall}}{\text{total number of words recalled}} \times 100$$

A 3 (Groups) X 2 (Instructional Conditions) X 2 (Order of Conditions) analysis of variance was performed on the cumulative recall scores. As in the clustering analysis, only the Group main effect was significant ($F = 3.73$, 2/24 df, $p < .05$). Referral to Table 3 reveals that the cumulative recall scores of first grade subjects were considerably lower than the scores of the other two groups.

The final analysis performed on recall scores was a 3(Groups) X 2(Instructional Conditions) X 2(Order of Conditions) X 9(Serial Position) analysis of variance. The only significant effect revealed by this analysis was the main effect for Serial Position ($F=6.27$; 8/192 df, $p < .001$). Two effects which approached significance, however, were the main effect for Group ($F = 2.91$, 2/24 df, $p < .10$), and the Instructional Condition X Serial Position interaction ($F = 1.72$, 8/192 df, $p < .10$). Figure 1 represents serial position curves by group and instructional condition. The curves demonstrate a striking primacy effect and a much less pronounced recency effect for each group and both conditions.

Insert Figure 1
about here

Pupillary Data

Reliability of measurement of pupil size from individual film frames was determined by having two scoreers obtain measurements independently for the first five subjects. While neither scorer was naive to the general design and purpose of the experiment, neither scorer was aware of the instructional condition for the frames being measured. Based on a correlation of measurements for the first five subjects (5 subjects X 18 blocks of frames/subject = 90 measurements), inter-scorer agreement was +.72.

After the initial determination of measurement reliability, one scorer measured the film frames for the remaining subjects. In an additional attempt to insure reliability of measurements, assessments of intra-scorer consistency were obtained by having each scorer select one subject's pupillary data for the other scorer to measure a second time. Intra-scorer agreement was +.96 for scorer 1 and +.89 for scorer 2.

Mean pupil size over the 20 frames immediately preceding the presentation of the first word constituted a baseline score for each subject. Baseline pupil size means by group and condition are presented in Table 4.

Insert Table 4 Here

Differences between means as a function of group and condition were assessed with t-tests. No significant differences were revealed when baseline pupil size was compared for the two instructional conditions within groups. Since no differences within groups as a function of condition existed, these means were collapsed across condition, and the overall group means were compared. The overall first-grade mean ($\bar{X}$ = 2.93, SD = .23) was significantly greater than both the overall fifth-grade mean ($\bar{X}$ = 2.71, SD = .10) (t = 4.40, 38 df, $p < .001$) and the overall MR mean ($\bar{X}$ = 2.67, SD = .42) (t = 2.36, 38 df, $p < .05$). The fifth-grade mean and the MR mean did not differ significantly.

The initial pupil size analysis was performed on deviation scores over the 18 blocks of frames in each condition. Neither of the predicted interactions involving groups and conditions were significant in this analysis. All variances, however, were found to be widely discrepant, and the assumption of homogeneity of variance did not seem warranted (F_{max} = 84.30, 9/108 df).

Natural logarithm ($\log_e$) transformations were performed on the pupillary measures. This procedure was justified in light of the finding with respect to heterogeneity of variance. Previous investigators have treated pupillary data in a similar fashion (Boersma, Wilton, Barham, & Muir, 1970). Since in some cases pupillary deviation scores for blocks of five frames were negative, a constant of 50 was added to each measurement to enable the transformations to be made.

Insert Figure 2 about here

Figure 2 shows changes in log dilation over the course of word presentation for the three groups. A 3(Groups) X 2(Instructional Condition) X 2(Order of Conditions) X 18(Blocks of 5 Frames) analysis of variance was performed on transformed pupil size scores. Significant effects were found for the Blocks of Frames main effect

($F=2.87$, 17/408 df, $p < .001$), the Groups X Blocks of Frames interaction ($F = 2.11$, 34/408 df, $p < .001$), and the Groups X Instructional Conditions X Blocks of Frames interaction ($F = 1.55$, 24/408 df, $p < .05$). In addition, both the Instructional Condition main effect ($F = 3.83$, 1/24 df, $p < .10$) and the Groups X Instructional Condition interaction ($F = 2.48$, 2/24 df, $p < .10$) approached significance.

It is apparent that the patterns of pupillary size over the course of word presentation were different for the three groups. For the first grade group, a sharp initial dilation was followed by a sizeable decrease in pupil size after the presentation of three or four words. The decline is greater in the Remember condition, as pupil size decreases at one point to a level lower than pre-presentation baseline. The fifth grade group's pupil size pattern is quite different. Although some variability is evident, a fairly consistent level of pupil size is maintained over the course of word presentation in both conditions. The pattern of pupil size of the retarded group is similar in some respects to that of the first-grade group, in that an initial dilation is followed by a decrease in pupil size. The two patterns differ, however, in that the rate of dilation is considerably slower for the retarded group, and the level does not begin to decline until around six words (12 blocks of frames) have been presented.

All subjects showed an initial dilation when word presentation began. Table 5 presents means by groups and conditions for magnitude of first dilation "peak" (i.e., the point at which pupil size first ceased to increase). Mean latencies, or times to first dilation peaks are also presented.

A 3 (Groups) X 2(Instructional Conditions) X 2(Order of Conditions) analysis of variance based on magnitude of first dilation revealed a significant main effect for Group ($F=11.16$, 2/24 df, $p < .001$) and a significant Group X Instructional Condition interaction ($F=7.27$, 2/24 df, $p < .01$). Inspection of the mean log dilation scores presented in Table 5 shows that while degree of dilation was greatest for the first grade group, the difference in dilation as a function of instructional condition was greatest for the fifth grade group.

A 3 (Groups) X 2(Instructional Conditions) X 2(Order of Conditions) analysis of variance was performed on latency, or time to first dilation peak. The only significant effects revealed by this analysis were the main effect for Instructional Condition ($F=4.49$, 1/24 df, $p < .05$) and the Order of Conditions X Instructional Conditions interaction ($F=7.02$, 1/24 df, $p < .05$).

Since large error terms in the original analysis for pupil size differences suggested substantial between-groups variability in pupil size, it was decided to determine if any group differences in variability existed. Standard deviations of untransformed (but with constant added) pupil deviation scores (18 per subject) were calculated for each subject. The average standard deviations for the three groups were: first grade, 11.14; fifth grade, 7.74; retarded, 6.26. Analyses of these differences were assessed by t-tests which revealed significant differences in variability between first grade and fifth grade subjects ($t = 2.48$, 8 df, $p < .05$) and between first grade and retarded subjects ($t = 2.65$, 8 df, $p < .05$), but no significance difference between fifth grade and revealed significant differences in variability between first grade and fifth grade subjects

($t = 2.48$, 8 df, $p < .05$) and between first grade and retarded subjects ($t = 2.65$, 8 df, $p < .05$), but no significant difference between fifth grade and retarded subjects.

Discussion

It must be concluded that these results support the hypothesis that deliberate memorization only gradually emerges as a type of encounter with information which is separate from perception (Appel, Cooper, McCarrell, Sims-Knight, Yussen, & Flavell, 1972). Younger children and retarded children remembered a comparable number of words under instructions to remember and instructions to listen for clarity. Older children, on the other hand, remembered significantly more words in the remember condition than in the listen condition.

Consideration of clustering, cumulative recall, and serial position data suggests some possible explanations of the differential effects of instructional conditions for the three groups. While there was more clustering in recall by the first grade group than for any other group, fifth grade subjects' mean cumulative recall index was greater than that for younger subjects. This finding indicates that the older children were possibly using cumulative rehearsal strategies during acquisition to a much greater degree than were younger children.

While the finding of less clustering with age is in conflict with the majority of developmental studies of clustering, several studies have reported failures to find increased clustering with age (Rossi & Rossi, 1965; Rossi & Wittrock, 1967; Schultz, Charness, & Berman, 1973). At least part of the discrepant findings with respect to clustering may be due to the use of different indices of clustering used by different investigators. Bousfield's clustering index (Cohen & Bousfield, 1956) is an index which is used widely, but it is perhaps inappropriate in developmental studies. The use of children as subjects in free recall tasks usually requires a short list with few categories and few items per category, but Bousfield's measure underestimates amount of clustering when few categories and few items per category are used.

Another possible source of the discrepancies in the literature concerns the way in which items are presented to the subjects. Successive versus simultaneous presentation may elicit different acquisitional strategies, since in one case (simultaneous) the subject has access to the entire set of stimuli from the beginning, and in the other case (successive) he does not. In the Appel et al. (1972) study, two experiments were reported, one involving successive and the other involving simultaneous presentation. For the simultaneous task, older children were found to cluster much more than younger children. On the successive task, the finding was that the clustering scores of older and younger groups were more similar. The present finding of <u>more</u> clustering by younger subjects could be due to the fact that the clustering index used in this study provides a more liberal estimate of clustering than does the index used by Appel et al. (1972) when the number of words recalled is relatively few.

The pupillary data, like the recall data, seem to support the differentiation hypothesis. As indicated by the significant interaction of Groups X Conditions X Blocks of Frames, the groups' pupil size patterns were different over the course of word presentation for the two instructional conditions. Likewise, the differentiation hypothesis is supported by the significant Groups X Conditions interaction found in the analysis performed upon first dilation magnitudes.

In the first grade group, there was a tendency for pupil size in the Remember condition to fall below that in the Listen condition after the presentation of three or four words. A similar rapid decrease in pupil size has been reported when adults were presented tasks which taxed their ability to process the information presented (Peavler, 1974). The hypothesis suggested by these results is that first grade subjects may have found the task too demanding or confusing after the presentation of several words, and subsequently suspended whatever efforts they were directing toward acquiring the information. The same phenomenon may have occurred in the retarded group, since their mean pupil size also decreased, but after the presentation of around six, rather than three or four words.

Even though the pupillary results lend some support to the differentiation hypothesis, they do not clarify the problem of knowing precisely what kinds of acquisitional strategies subjects were using. Kahneman has suggested that in order to make specific inferences from pupillary data, the investigator must have a theory of what the subject is doing during the course of the task (Kahneman & Wright, 1971). In other words, subjects may engage in different cognitive activities during almost any particular task, and the pupillary response reflects this ambiguity wherever it exists. In the present study, some independent evidence suggests that subjects were using at least two different strategies (cumulative recall and conceptual clustering). It may be most accurate in this case to maintain that patterns of pupil size were reflective of cognitive effort in general rather than of particular cognitive strategies. This interpretation is one which can incorporate the failure to discover any significant interactions involving instructional condition for the clustering and cumulative recall measures. Instructional conditions may have had their greatest effect on motivation, or the propensity to exert directed effort towards remembering the lsit and a lesser effect on the particular type of acquisitional strategy used.

With respect to the contribution of the results of the present study to the understanding of nature of memory processes in retarded persons, several points can be made. The recall performance data indicate that the retarded group's recall was more similar to that of the first grade group than to that of the fifth grade group -- their CA peers. Inspection of the serial position curves, and the clustering and cumulative recall scores, however, indicates more similarity between the retarded and the fifth grade groups than between the retarded and the first grade groups. This result suggests that while retarded subjects may have been using acquisitional strategies similar to those used by their CA peers, their use of those strategies may have been less efficient. The retarded subjects, then, may have chosen appropriate acquisitional strategies, but were unable to apply them as efficiently as their CA peers, and were unable to differentiate their use of the strategies as a function of task demands. Fifth grade subjects, on the other hand, were able to both choose appropriate strategies and to apply them efficiently in accord with the requirements of the two tasks.

The measures of variability of pupil size during word presentation may be indicative of the consistency of effort distributed over the span of presentation. Variability was relatively great for the first grade group, and this inconsistency may be reflective of their inability to generate sustained effort in the form of application of acquisitional strategies. The fifth grade and retarded groups showed less variability, suggesting that they were more able to exert directed sustained effort, although fifth grade nonretarded subjects' efforts resulted in more efficient use of the acquisitional strategies chosen.

Table 1

CA and IQ Characteristics of Experimental Groups

		CA (yr.-mo.)		IQ	
Group	N	Mean	SD	Mean	SD
1st Grade	10	6- 7	2.12		
5th Grade	10	10-11	5.87		
EMR (5th Grade)	10	11-10	8.85	67	7.72

Table 2

Number of Words Recalled

	Restricted Sample (N=30)					
	1st Grade (n=10)		5th Grade (n=10)		MR (n=10)	
	Remember	Listen	Remember	Listen	Remember	Listen
Mean	3.40	2.70	5.50	3.40	3.80	3.90
SD	1.17	1.57	1.18	2.32	1.69	1.45
	Complete Sample (N=47)					
	1st Grade (n=19)		5th Grade (n=13)		MR (n=15)	
	Remember	Listen	Remember	Listen	Remember	Listen
Mean	3.74	2.84	5.54	3.38	3.53	3.67
SD	1.14	1.54	1.05	2.10	1.46	1.40

Table 3

Clustering and Cumulative Recall Scores

		Group		
		1st Grade	5th Grade	MR
Clustering	$\bar{X}$	58.20	23.60	38.30
	SD	64.06	36.21	35.01
Cumulative Recall	$\bar{X}$	5.80	25.45	24.45
	SD	14.51	28.02	23.30

Table 4

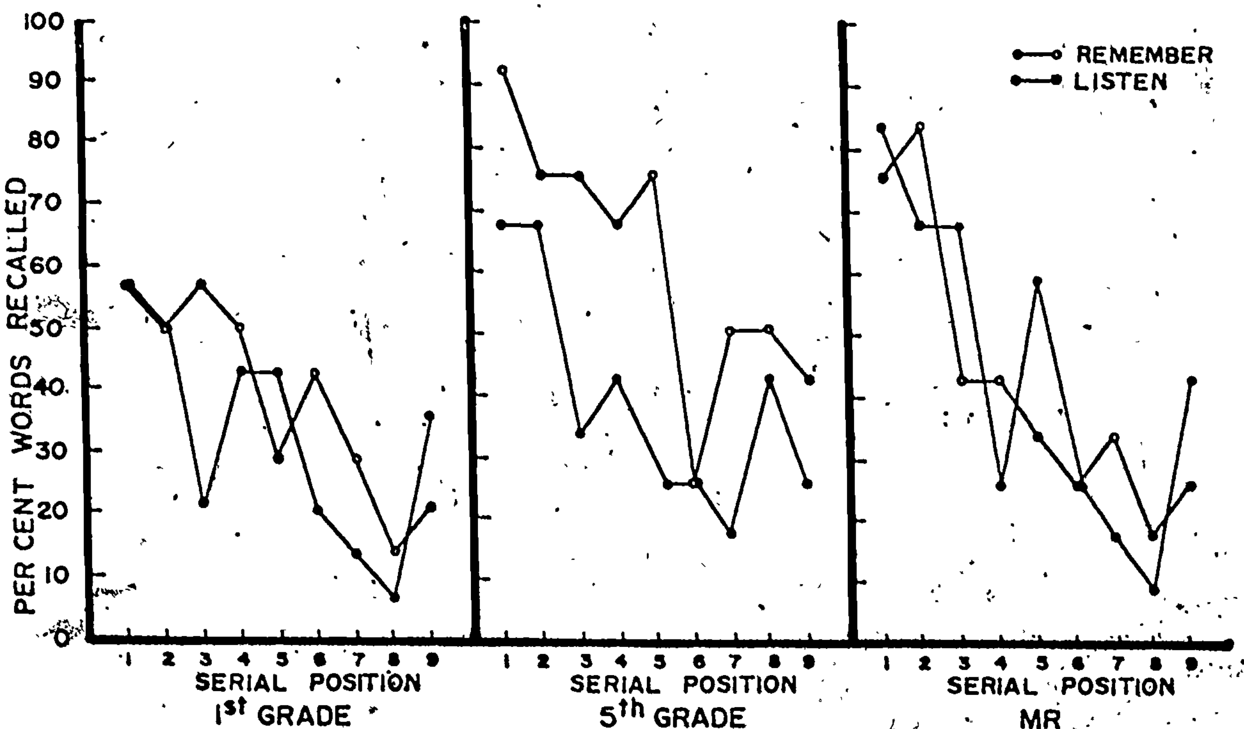
Baseline Log Pupil Size

	1st Grade		5th Grade		MR	
	Remember	Listen	Remember	Listen	Remember	Listen
Mean	2.90	2.97	2.70	2.72	2.60	2.75
SD	.31	.13	.11	.11	.44	.39

Table 5

Magnitude and Latency of First Dilation

		Group					
		1st Grade		5th Grade		MR	
		R	L	R	L	R	L
Magnitude, 1st Dilation (log_e)	$\bar{X}$	2.90	3.25	2.87	2.22	1.99	2.25
	SD	.65	.24	.42	.74	.52	.50
Latency, 1st Dilation (seconds)	$\bar{X}$	9.50	9.00	10.00	22.00	10.50	15.00
	SD	4.38	4.59	7.82	18.29	6.85	17.00



REMEMBER
LISTEN
PER CENT WORDS RECALLED
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9
SERIAL POSITION
1st GRADE
SERIAL POSITION
5th GRADE
SERIAL POSITION
MR

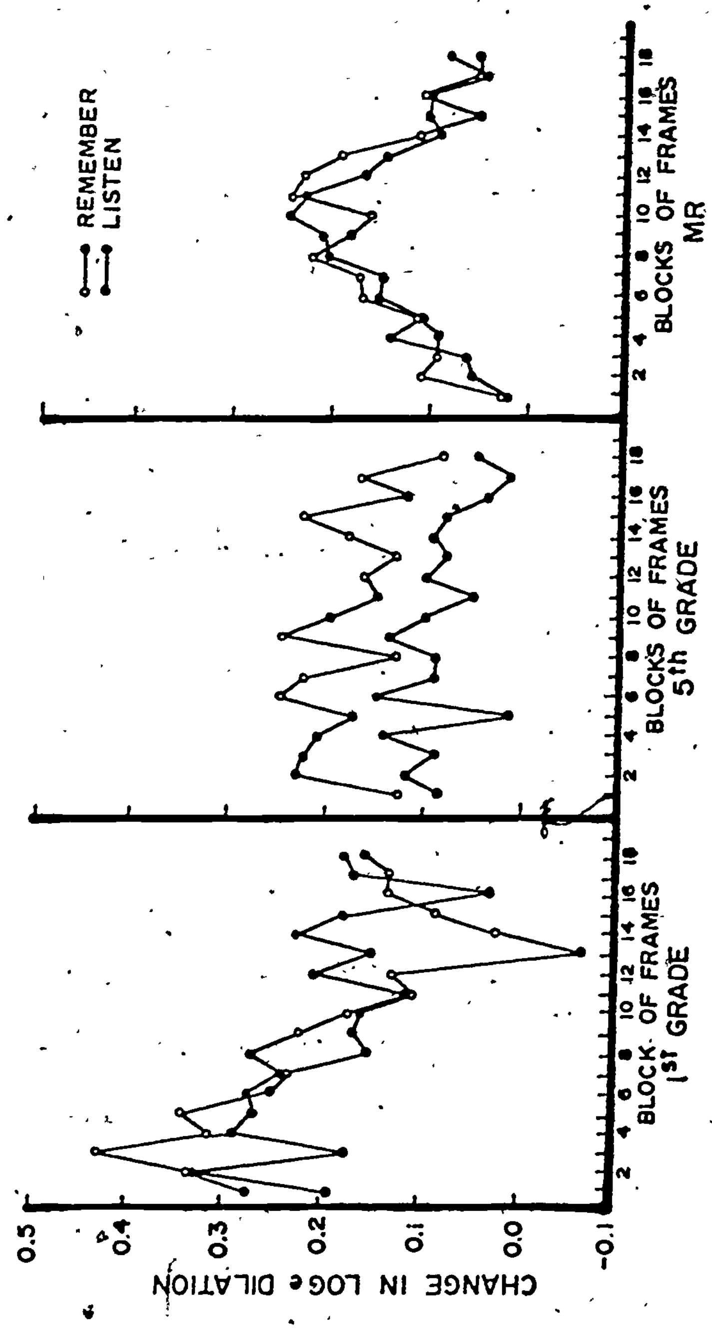

REMEMBER
LISTEN
CHANGE IN LOGe DILATION
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
2 4 6 8 10 12 14 16 18
BLOCK OF FRAMES
1st GRADE
BLOCKS OF FRAMES
5th GRADE
BLOCKS OF FRAMES
MR

Figure Captions

Figure 1. Serial position curves.

Figure 2. Change in log pupil size by group and condition.

Footnotes

[1]This clustering index is described by Frankel and Cole (1971) as a percentage in clusters (PC) measure, and was selected because it allowed for maximal comparison with a similar index of cumulative recall.

Experiment II: Pupil Size Changes of Retarded and Nonretarded Adolescents During Digit Memory Tasks.

Experiment II was designed to assess the sensitivity of pupil size measurement to differences in processing of digits. Digit memory has been frequently studied in retarded populations (e.g. Ellis, 1970), and poor digit memory is almost always found to covary with low IQ. According to Ellis (1970) interpretation of poor digit recall by retarded subjects, deficiencies in digit recall performance are due to failures to rehearse digits actively as they are presented. If pupil size changes are reflective of active cognitive processing, patterns of pupil size changes should be different for subjects who perform well on digit recall tasks and those who perform poorly.

An additional aspect of Experiment II concerned an attempt to replicate one of the outcomes of Experiment I. Some evidence for differentiation of processing between the Listen and Remember conditions was found in Experiment I for the older subjects. Administration of digit memory tasks under Listen and Remember conditions in the present experiment provided a further test of the differentiation hypothesis. Both retarded and nonretarded groups were expected to differentiate their processing efforts as a function of pre-task instruction. In order to increase the validity of the Listen instruction, three consecutive trials were administered in that condition, followed by three trials with a Remember instruction.

Procedure

Subjects listened to taped six-digit strings (one digit every two seconds) under two instructional conditions, an instruction only to listen to the digits (L condition), and an explicit instruction to try to remember the digits exactly as they were heard (R condition). Each subject first listened to three consecutive digit strings under the L condition. Instructions to a subject before trial L_1 were as follows: "You are going to hear some numbers. I want you to listen to the numbers, and after you listen, I want you to tell me if you could hear them clearly. Do you understand? O.K., lets begin." Before trials L_2 and L_3, the experimenter said, "Here are some more numbers. See if you can hear these numbers clearly." Following trial L_3, subjects received an unexpected recall test. The experimenter said, "I know I didn't ask you to try to remember what numbers you heard, but can you remember the numbers you just heard? Try to remember then exactly as you heard them." The experimenter then recorded the subject's report of recalled digits. Before the beginning of trial R_1, subjects were given the following instruction: "I'd like for you to listen to some more numbers. This time, after you hear each group, I will ask you to tell me exactly what numbers you heard. Try to remember them just as you hear them. Do you understand? O.K., lets begin." Following Trials R_1, R_2, and R_3, subjects' recall of digits were recorded.

A procedure for scoring digit recall was developed which gave credit for digits correctly recalled as well as credit for order of recall. The scoring criteria were as follows: 1 point for each digit correctly recalled (six points possible), one point for each pair of digits in correct order (five points possible), and one point for each initial or terminal digit in proper position (two points possible). Maximum digit recall score possible, then, was 13.

During all digit presentation trials, pupil size was photographically monitored. The camera ran continuously for 30 seconds (two frames/second) on each trial. Three separate phases of each trial were identified: (1) Pre-Digits Baseline Phase: 10 seconds immediately preceding the first digit, (2) Digits Phase: 10 seconds during digit presentation, and (3) Post-Digits Phase: 10 seconds immediately following presentation of the sixth digit in each string.

Subjects

Ten students from an EMR classroom and eleven students from a regular classroom in the same high school participated in the study. IQ scores for the EMR students was obtained from their records, and each score was based on an administration of either the Stanford-Binet or WISC not more than two years prior to the date of the present experiment. Mean IQ for the retarded group was 68.60 (SD=11.46). Subjects from the regular classroom were administered the Peabody Picture Vocabulary Test, and a mean IQ of 90.90 (SD=11.57) resulted. The two groups were matched for CA. Mean CA for the retarded group was 219.00 mo. (SD=12.61 mo.), and mean CA for the nonretarded group was 217.09 mo. (SD=7.90 mo.). The groups were not equally constituted by sex, as the sample was comprised of 14 boys and 7 girls.

Results

Digit recall performance for the two groups is illustrated

Insert Figure One
About Here

in Figure 1. A 2(Groups) X 4 (Trials) analysis of variance performed on these data resulted in significant main effects for Groups (F=45.36, 3/57 df, $p < .001$), and a significant Groups X Conditions interaction (F = 7.04, 3/57 df, $p < .001$). For both retarded and nonretarded groups, Tukey multiple comparisons indicated a significant ($p < .01$) increase in recall performance on the first trial under the intentional memory instruction (R_1). When additional comparisons were made between performance on intentional memory trials (R_2; R_3) and performance on the incidental trial (L_3), it was found that by trial R_3, recall by retarded subjects was not significantly better than their recall on trial L_3. As is indicated in Figure 1, performance by nonretarded subjects gradually improved across intentional memory trials, while that of retarded subjects declined. A further breakdown of the significant Groups X Conditions interaction was performed by conducting Tukey comparisons between groups on each trial. Those analyses showed that only on Trial R_3 did the difference in performance between groups reach significance ($p < .01$).

Pupil size data were analyzed for eight subjects in each group. Data for a total of five subjects had to be excluded due to an excessive number of unscoreable frames.

The procedure of averaging pupillary deviations from baseline over five frames to constitute blocks of frames (BFs) described in the general procedures section was followed in the present experiment. For every trial, each subject's data consisted of an individual baseline mean pupil size (averaged over the first four BFs), four BFs during digit presentations (digits phase), and four BFs immediately following presentation of the last digit (post-digits phase).

The first analysis performed on the pupillary data was a 2 (Groups) X 2 (Conditions) X 3(Trials) X 2 (Phase) X 4 (BFs) split-plot factorial analysis of variance with the Groups factor as the sole between-subjects factor (Kirk, 1968). Significant effects were revealed only for two higher-order interactions, the four-way interaction involving Groups, Trials, Phase, and BFs (F = 3.00, 6/84 df, $p < .05$) and the five-way interaction involving all factors (F = 2.35, 6/84 df, $p < .05$). To simplify interpretation of these complex interactions, separate analyses of variance were performed on pupil size data for the Digits and Post-Digits Phases broken down by individual trials. For Trial 1 data (L_1 vs R_1) during the Digit Phase, the only significant effect was the Conditions X BFs interaction (F=5.74, 3/42 df, $p < .005$). Figure 2 suggests that this

---

Insert Figure 2
about here

---

interaction is due to gradually increasing pupil size for both groups in the Remember Condition over trials, accompanied by a general decrease for both groups in the Listen Condition. No significant effects were found for Trial 2 data (L_2 vs R_2). For Trial 3 data (L_3 vs R_3), a significant main effect for Groups (F = 8.71, 1/14 df, $p < .05$) during the Post-Digits Phase indicated greater pupil size for the nonretarded group.

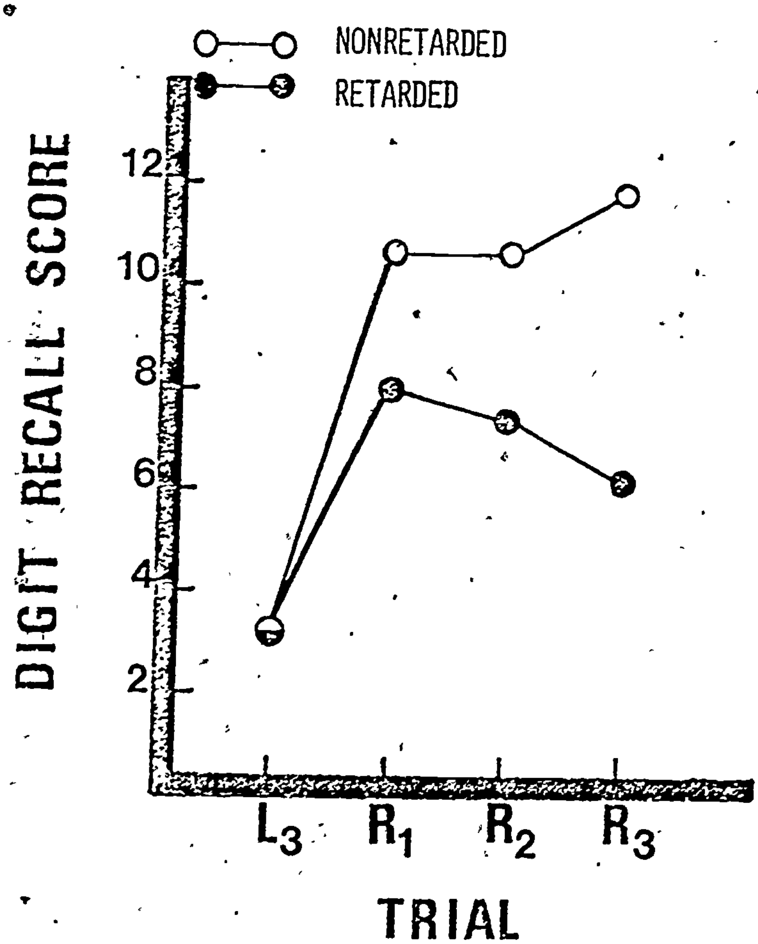
Since the "Listen" instruction should have attained maximum credibility by Trial L_3, a comparison of pupil size changes between Trials L_3 and R_1 was of considerable interest. For that comparison, a significant Conditions X Bfs interaction was found for both the Digits ($F=4.60$, 3/42 df, $p < .01$) and the Post-Digits Phase ($F = 2.83$, 3/42 df, $p < .05$).

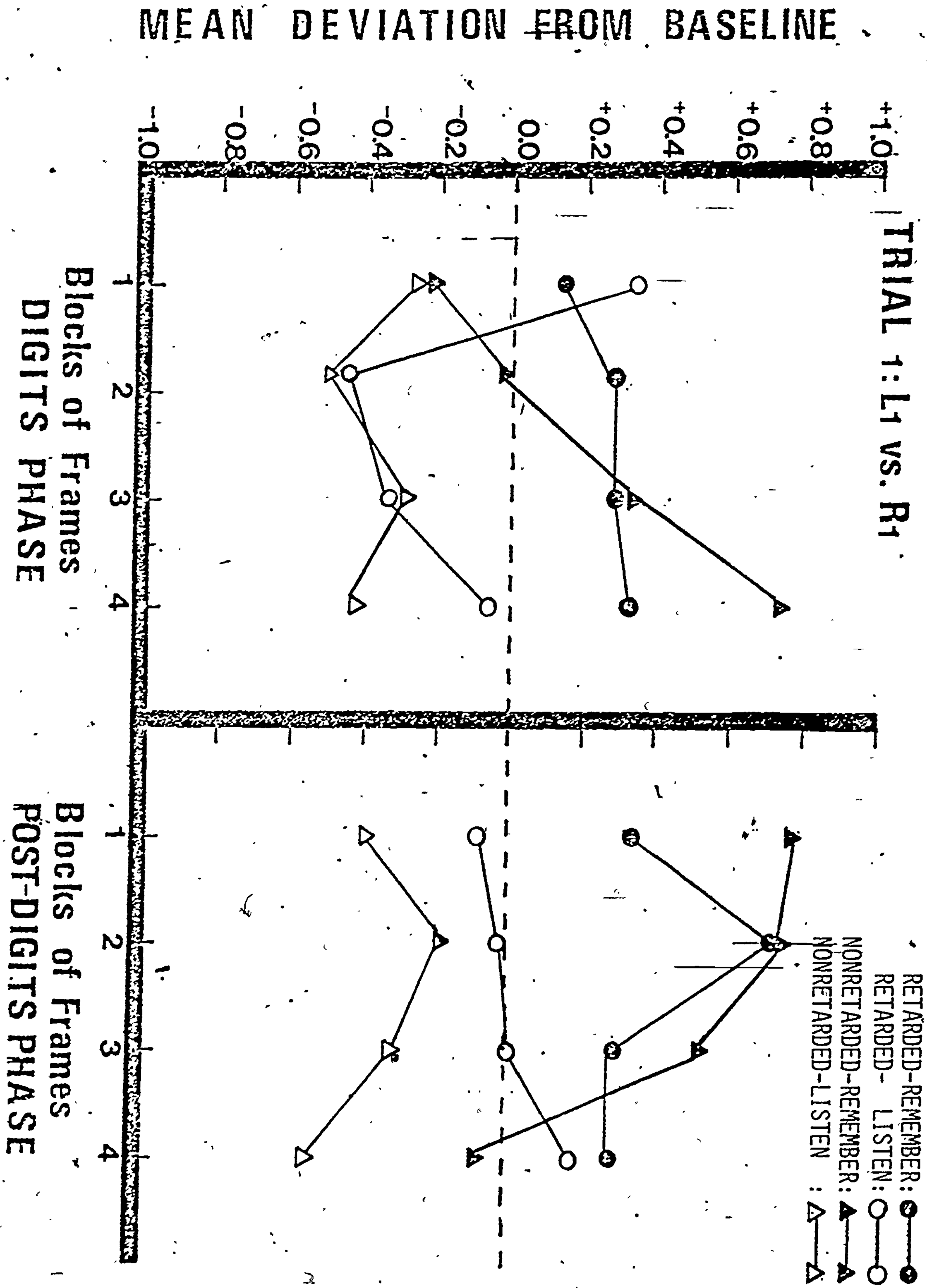
Discussion

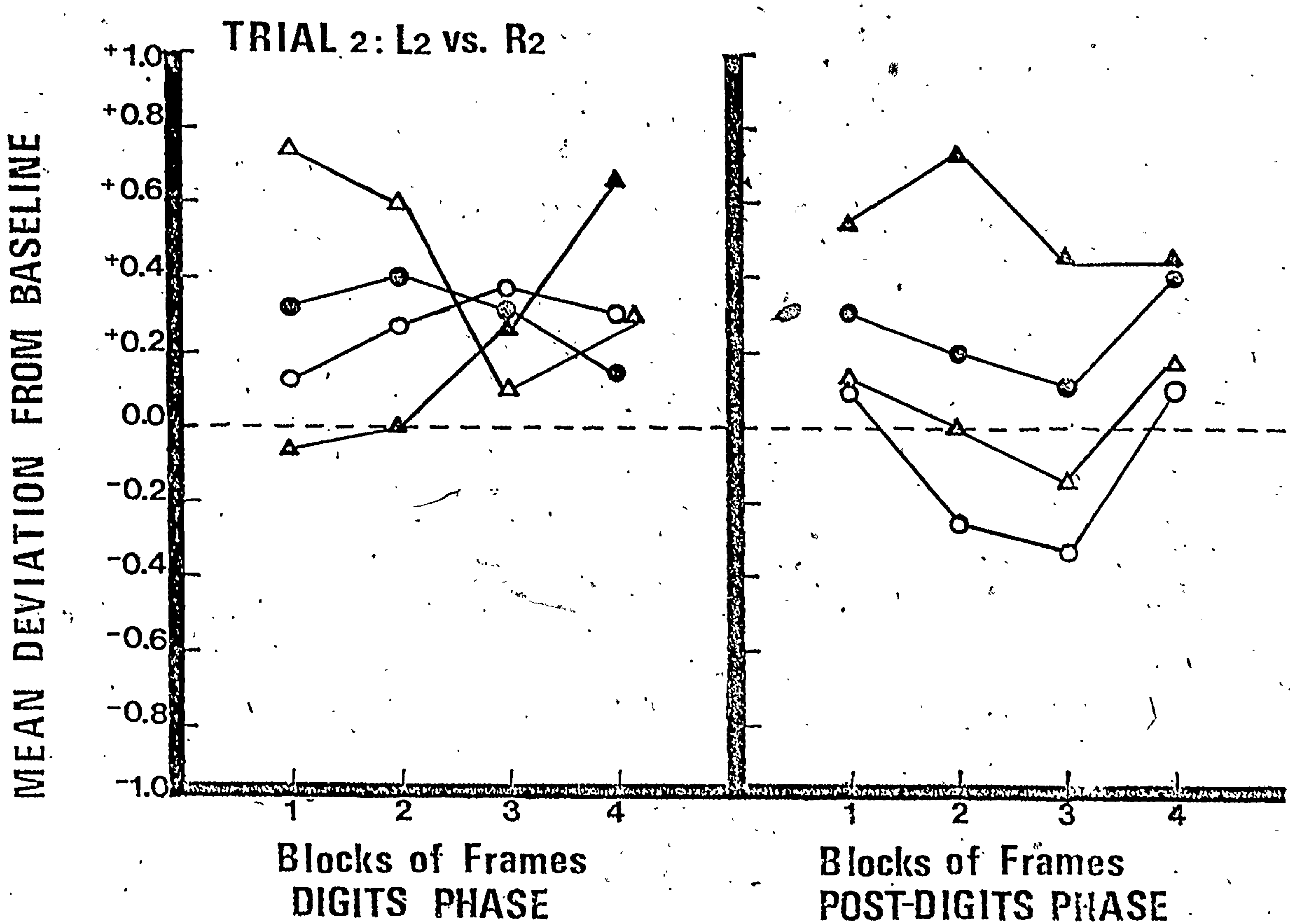
Considering the digit recall data first, it is evident that while both groups of subjects demonstrated differentiation of effort for the two instructional conditions, Group differences are evident for performance of the task across trials. While nonretarded subjects' performance was found to improve over the intentional memory trials, the performance of retarded subjects declined over those trials to the point that their performance was significantly poorer by Trial R_3. Since one would ordinarily expect improvement over trials on the same type of task as a result of practice, we might suppose that the retarded subjects were losing interest in the task and investing less and less cognitive effort in trying to rehearse the digits actively during presentation.

With regard to pupillary data, the result of most interest is that th only significant effect involving the Groups factor was found in the Trial 3 analysis. That is, the significant Groups difference on Trial R_3 obtained in the analysis of digit recall performance appears to have been paralleled by a similar Groups difference when pupil size data were analyzed.

The pupillary data were expected to reflect differences as a function of instructional condition, since such differences were found for the digit recall data. While some evidence was found to support that expectation (significant Conditions X BFs interactions in three of the analyses), pupil size differences as a function of instructional condition were not found consistently across trials.


NONRETARDED
RETARDED
DIGIT RECALL SCORE
12
10
8
6
4
2
L_3
R_1
R_2
R_3
TRIAL


MEAN DEVIATION FROM BASELINE
TRIAL 1: L1 vs. R1
+1.0
+0.8
+0.6
+0.4
+0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1
2
3
4
Blocks of Frames
DIGITS PHASE
Blocks of Frames
POST-DIGITS PHASE
RETARDED-REMEMBER:
RETARDED- LISTEN:
NONRETARDED-REMEMBER:
NONRETARDED-LISTEN :



TRIAL 2: L2 vs. R2
MEAN DEVIATION FROM BASELINE
+1.0
+0.8
+0.6
+0.4
+0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1
2
3
4
Blocks of Frames
DIGITS PHASE
1
2
3
4
Blocks of Frames
POST-DIGITS PHASE

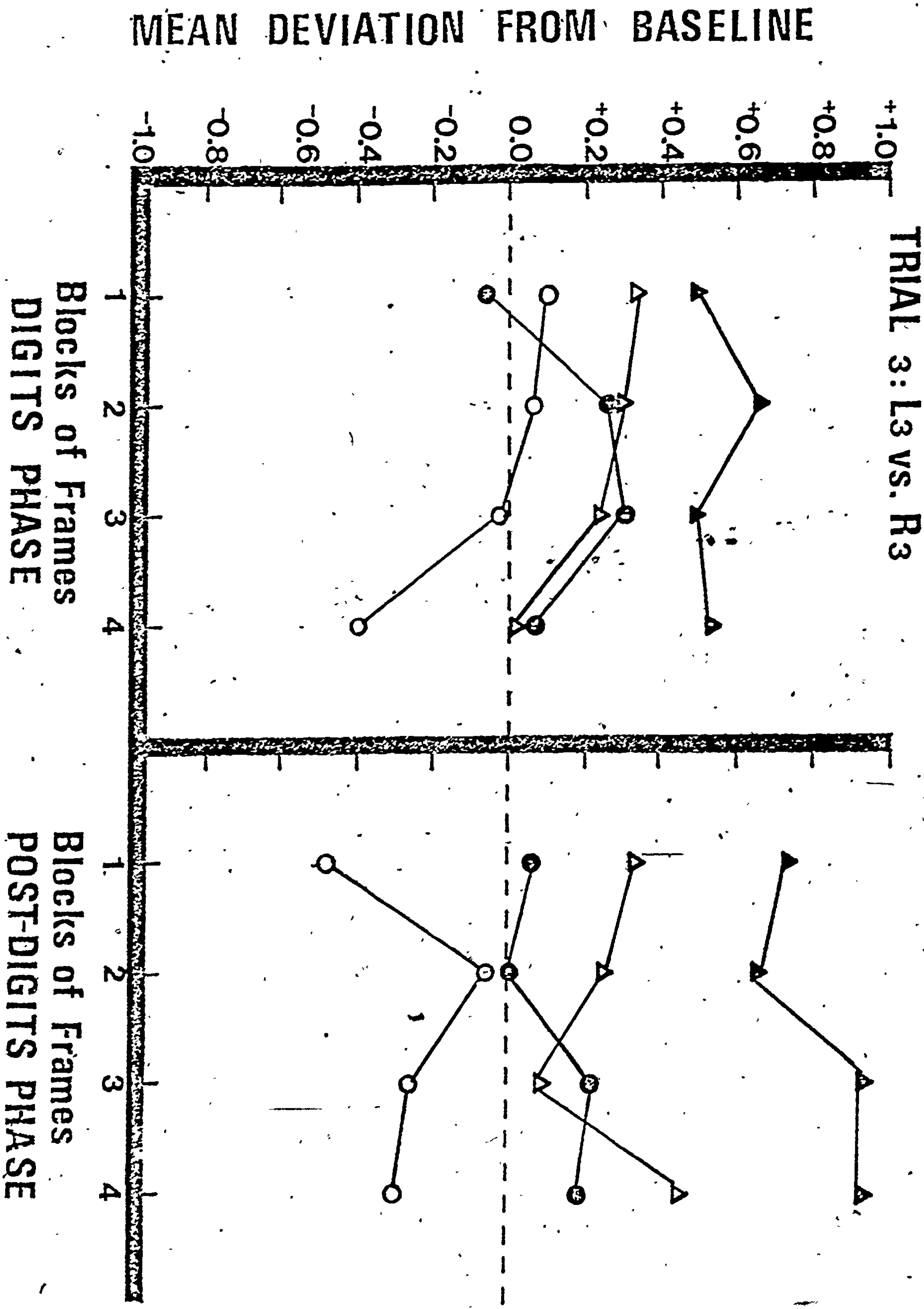

MEAN DEVIATION FROM BASELINE
TRIAL 3: L3 vs. R3
+1.0
+0.8
+0.6
+0.4
+0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1
2
3
4
Blocks of Frames
DIGITS PHASE
Blocks of Frames
POST-DIGITS PHASE

Figure Captions

Figure 1 Digit recall scores

Figure 2 Change in pupil size

General Conclusions

The primary purpose of these studies was to determine the effectiveness of pupillary measurements as indicators of covert cognitive processes in retarded and non-retarded children. Since so little research of a similar nature had been done with children as subjects, it was decided to collect pupillary data while children performed memory tasks which are fairly common in developmental research. A necessary condition for making conclusions about the worth of pupillary measures was that the memory results conform to a large degree to those results obtained by previous investigators.

In both Experiments I and II, this necessary condition was satisfied. The confirmation in Experiment I was that older children remembered more words given an instruction to remember them than they did when given an instruction merely to listen to the words. Younger children and retarded childrens' recall did not differ in the two conditions. Thus the major result obtained by Appel et.al. (1972) was replicated. In Experiment II, the confirmation was that intentional digit memory was greater for nonretarded than for retarded subjects, a result obtained in a great number of investigations (e.g. Ellis, 1970).

With these necessary conditions satisfied, the assumption was that relatively unambiguous interpretation of pupillary measure could be made. Since much research indicates that developmental memory differences are due to differential use of active processing strategies (e.g. Flavell, 1970), and since the results of many studies done with adults indicate that pupillary measures can reflect cognitive activity (e.g. Goldwater, 1972), differences in pupillary measures were expected to parallel memory differences found in the present experiments.

While the pupillary data are suggestive in several respects, a close correspondence with recall data was obtained in neither experiment. A major problem may have been the relatively low power of the statistical techniques employed. Cumbersome and time-consuming data collection procedures limited the number of subjects' data available for the anlyses performed on pupillary data. Since intra- and intersubject variability in pupil size changes was so great, differences due to experimental conditions or group membership may have been obscured by the small Ns.

Because the literature indicates that pupil size measurements have been of some use on cognitive tasks with adult subjects, and the majority of those studies have involved relatively few subjects, the failure to obtain unambiguous results in the present studies is puzzling. One possible explanation is that low-CA and low-MA subjects may exhibit greater variability on physiological measures than do nonretarded adult subjects. Baumeister (1974) has proposed that the performance of retarded subjects on a great number of different tasks is characterized by a degree of intra- and inter-subject variability greater than that found in nonretarded populations.

An alternative explanation of the lack of clarity in the pupillary findings is that uncontrolled motivational factors confounded the results. Since pupil size has been demonstrated to reflect states of arousal (Goldwater, 1972), different emotional reactions to the experimenter or to various aspects of the experimental tasks could have affected the results.

Whatever the explanation, however, two conclusions seem clear. The first is that more research is needed if a definitive answer to the questions posed in the present studies are to be answered. The research should unquestionably involve greater numbers of subjects and more sophisticated instrumentation to aid in data collection.

The second conclusion may conflict somewhat with the first. It is that investigators will have to weigh the potential contribution of data such as these against their cost in terms of time and resources invested. If questions which are basic to our knowledge of developmental differences in cognitive processing ability can be answered through the use of techniques which are simpler and more straightforward than the collection of pupillary data, then perhaps those alternatives should be exercised. If, on the other hand, techniques for collection of pupillary data can be refined somewhat, and if the potential information obtained is of great importance and either inaccessible or more difficult to obtain through alternative measures, the collection of such data should be encouraged.

REFERENCES

Adams, J.A. Human Memory. New York: McGraw-Hill, 1967

Appel, L.F., Cooper, R.G., McCarrell, N., Sims-Knight, J., Yussen, S.R., & Flavell, J.H. The development of the distinction between perceiving and memorizing. Child Development, 1972, 43, 1365-1381.

Baumeister, A.A. Personal Communication, 1974

Belmont, J.M. Long-term memory in mental retardation, In N. R. Ellis (Ed.) International Review of Research in Mental Retardation. New York: Academic Press, 1966, 1, 219-255.

Belmont, J.M. and Butterfield, E.C. The relation of short-term memory to development and intelligence. In L.P. Lipsett and H.W. Reese (Eds.) Advances in Child Development and Behavior, New York: Academic Press, 1969, 4.

Belmont, J.M., and Butterfield, E.C. What the development of short-term memory is. Human Development, 1971, 14, 236-248. (a)

Belmont, J.M., and Butterfield, E.C. Learning strategies as determinants of memory deficiencies. Cognitive Psychology, 1971, 2, 411-420. (b)

Boersma, F., Wilton, K., Barham, R., and Muir, W. Effects of arithmetic problem difficulty on pupillary dilation in normals and educable retardates. Journal of Experimental Child Psychology, 1970, 9, 142-155.

Bradshaw, J.L. Pupil size as a measure of arousal during information processing. Nature, 1967, 216, 515-516.

Bradshaw, J.L. Pupil size and problem solving. Quarterly Journal of Experimental Psychology, 1968, 20, 116-122.

Cohen, B.H., & Bousfield, W.A. The effects of a dual-level stimulus-word list on the occurrence of clustering in recall. Journal of General Psychology, 1956, 55, 51-58.

Coleman, F.D., & Paivio, A. Pupillary dilation and mediation processes during paired-associate learning. Canadian Journal of Psychology, 1970,24,261-270.

Daehler, M.W., Horowitz, A.B., Wynns, F.C., & Flavell, J.H. Verbal and nonverbal rehearsal in children's recall. Child Development, 1969, 40, 443-452.

Ellis, N.R. Memory processes in retardates and normals: Theoretical and empirical considerations. In N.R. Ellis (Ed.), International Review of Research in Mental Retardation. Vol. 4. New York: Academic Press, 1970.

Flavell, J.H. Developmental Studies of mediated memory. In H.W. Reese & L.P. Lipsett (Eds.), Advances in Child Development and Behavior. Vol. 5. New York: Academic Press, 1970.

Flavell, J.H., Beach, D.H., & Chinsky, J.M. Spontaneous verbal rehearsal in a memory task as a function of age. Child Development, 1966, 37, 283-299.

Flavell, J.H., Friedrich, A.G., and Hoyt, J.D. Developmental changes in memorization processes. Cognitive Psychology,1970, 1, 324-340.

Frankel, F., & Cole, M. Measures of Category clustering in free recall. Psychological Bulletin, 1971, 76, 39-44.

Goldwater, B. Psychological significance of pupillary movements. Psychological Bulletin, 1972, 77, 340-355.

Hagen, J.W. Some thoughts on how children learn to remember. Human Development, 1971, 14.

Hess, E.H. Attitude and pupil size. Scientific American, 1965, 212, 46-54.

Hess, E.H., & Polt, J.M. Pupil size as related to interest value of visual stimuli. Science, 1960, 132, 349-350.

Hess, E.H. & Polt, J.M. Pupil size in relation to mental activity during simple problem-solving. Science, 1964, 143, 1190-1192.

Hunt, J. McV. Intelligence and Experience. New York: Ronald Press, 1961.

Jensen, A.R. and Rohwer, W.D., Jr., The effect of verbal mediation on the learning and retention of paired-associates by retarded adults. American Journal of Mental Deficiency, 1963, 68, 80-84.

Johnson, D.A. Pupillary responses during a short-term memory task: Cognitive processing, arousal, or both? Journal of Experimental Psychology, 1971, 90, 311-318.

Kahneman, D. and Beatty, J. Pupil diameter and load on memory. Science, 1966, 154, 1583-1585.

Kahneman, D. and Beatty, J. Pupillary responses during a pitch-discrimination task. Perception and Psychophysics, 1967, 2 101-105.

Kahneman, D., Onuska, L., and Wolman, R.E. Effects of grouping on the pupillary response in a short-term memory task. Quarterly Journal of Experimental Psyhcology. 1968, 20, 309-311.

Kahneman, D., and Peavler, W.S. Incentive changes and pupillary changes in association learning. Journal of Experimental Psychology, 1969, 79, 312-318

Kahneman, D., and Wright, P. Changes in pupil size and rehearsal strategies in a short-term memory task. Quarterly Journal of Experimental Psychology, 1971, 23, 187-196.

Keeney, T.J., Cannizzo, S.R., & Flavell, J.H. Spontaneous and induced verbal rehearsal in a recall task. Child Development, 1967, 38, 953-966.

Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks-Cole, 1968.

Lowenfeld, I.E. Mechanisms of reflex dilation of the pupil. Historical review and experimental analysis. Documenta Opthamologica, 1958, 12, 185-448.

MacMillan, D.L. Facilitative effect of verbal mediation on paired-associate learning by EMR children. American Journal of Mental Deficiency, 1970, 74, 611-615.

Neimark, E., Slotnick, N.S., and Ulrich, T. Development of memorization strategies. Developmental Psychology, 1971, 5, 427-432.

Nunnally, J.C., Knott, P.D., Duchnowski, A., and Parker, R. Pupillary responses as a general measure of activation. Perception and Psychophysics, 1967, 2, 149-155.

Paivio, A. and Simpson, H.M. The effect of word abstractness and pleasantness on pupil size during an imagery task. Psychonomic Science, 1966, 5, 55-56.

Payne, D.T., Parry, M.E., and Harasymiw, J. Percentage of pupillary dilation as a measure of item difficulty. Perception and Psychophysics, 1968, 4, 139-143.

Peavler, W.S. Pupil size, information overload, and performance differences. Psychophysiology, 1974, 11, 559-566.

Rossi, E.L. & Rossi, S.I. Concept utilization, serial order and recall in nursery school children. Child Development, 1965, 36, 771-778.

Rossi, S.I. & Wittrock, M.C. Clustering versus serial ordering in recall by four year old children. Child Development, 1967, 38, 1139-1142.

Schaefer, T., Jr., Ferguson, J.B., Klein, J.A., & Rawson, E.B. Pupillary responses during mental activities. Psychonomic Science, 1968, 12, 137-138.

Schultz, T.R., Charness, M., & Berman, S. Effects of age, social class, and suggestion to cluster on free recall. Developmental Psychology, 1973, 8, 57-61.

Sokolov, E.N. Perception and the conditioned reflex. New York: Macmillan, 1963.

Spitz, H.H. Consolidating facts into the schematized learning and memory of educable retardates. Paper presented at Gatlinburg Conference on Research and Theory in Mental Retardation, Gatlinburg, Tennessee, March, 1972.

Spitz, H.H. The role of input organization in the learning and memory of mental retardates. In N.R. Ellis (Ed.) International Review of Research in Mental Retardation. New York: Academic Press, 1966, 2, 29-56.

Spitz, H.H. and Webreck, C.A. Effects of spontaneous vs. externally-cued learning on the permanent storage of a schema by retardates. American Journal of Mental Deficiency, 1972, 77, 163-168.

Stanners, R.F., and Headley, D.B. Pupil size and instructional set in recognition and recall. Psychophysiology, 1972, 9, 505-511.

Stanners, R.F. and Headley, D.B., & Clark, W.B. The pupillary response to sentences: Influences of listening set and deep structure. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 257-263.

Taylor, A.M., Josberger, M., and Knowlton, J.Q. Mental elaboration and learning in EMR children. American Journal of Mental Deficiency, 1972, 77, 69-76.

Turnure, J.E., and Walsh, M.K. Extended verbal mediation in the learning and reversal of paired-associates by educable mentally retarded children. American Journal of Mental Deficiency, 1971, 76, 60-67.

Wright, P., & Kahneman, D. Evidence for alternative strategies of sentence retention. Quarterly Journal of Experimental Psychology, 1971, 23, 197-213.