

STORY RECALL BY MENTALLY RETARDED CHILDREN¹

ELLEN H. BACON

AND

DAVID C. RUBIN

Duke University Medical Center

*Duke University*²

Summary.—Three stories were read to 19 mildly retarded young adolescents and to 19 nonretarded children of comparable mental age. The two groups did not differ consistently in the amount they recalled orally. Moreover, the two groups tended to recall the same parts of the stories, indicating that both groups are sensitive to the stories' structure.

From kindergarten to adulthood, developmental differences in story comprehension appear to be small. Although older subjects remember and elaborate more than young children, even kindergarten and first graders demonstrate sensitivity to the structure of stories (Brown & Smiley, 1977; Christie & Schumacher, 1975; Mandler & Johnson, 1977; Stein & Glenn, 1979). Likewise, retarded children seem to show little deficit in story recall. For instance, Paris, Mahoney, and Buckhalt (1974) used a semantic integration task to demonstrate that mildly retarded children comprehend and remember semantic relationships expressed in three connected sentences, suggesting that the mildly retarded children are capable of abstracting semantic relationships in a manner qualitatively similar to nonretarded children and adults. Similarly, Luftig and Johnson (1982) found that, while retarded children recalled less than nonretarded children of the same mental age, retarded children were able to identify important parts of the story and recalled the important parts more frequently than the unimportant parts.

The relative lack of differences in the way children of different age or intelligence process stories is not an isolated finding. Similar results have been found in a diverse set of populations. Graybeal (1981) categorized two stories using Mandler and Johnson's (1977) story grammar into Setting, Beginning, Simple Reactions—Goal, Attempt, Outcome, and Ending. She noted that, while language-impaired subjects recalled less than their control group, the two groups showed no difference in the relative proportion of times they recalled propositions from the six categories. Both groups were equally sensitive to the story structure. Rubin (1978) reported that junior high school, college, and elderly subjects tended to recall the same units from stories, and Rubin, Olson, Richter,

¹This research is an extension of the first author's doctoral dissertation supervised by Dr. Ann Turnbull. Her support and that of Dr. Bob Thompson, Director of the Duke Developmental Evaluation Clinic, are acknowledged. The authors thank Joyce Wasdell and the Durham County School system for their cooperation. The stories used divided into units as well as the scored data are on file with Microfiche Publications, POB 3513, Grand Central Station, New York, NY 10017. Request Document NAPS-04125, remit \$4.00 for fiche or \$9.85 for photocopy.

²Reprints are available from David C. Rubin, Department of Psychology, Duke University, Durham, North Carolina 27706.

and Butters (1981) noted that institutionalized schizophrenics and their controls as well as alcoholic Korsakoff patients and their controls tended to recall the same units from the Wechsler Memory Scale stories (Stone, Girdner, & Albrecht, 1946), although the standard differences in amount recalled were observed. Stories appear to be processed in similar ways by quite different clinical populations. That is, some populations who show marked cognitive deficits in given areas are nonetheless sensitive to story structure and show little difference from their controls in which units they are most likely to recall.

As demonstrated in the preceding paragraphs, to examine differences and similarities in the way stories are processed it is useful to ask not only how many units of a story are recalled but also which units are recalled. That is, in addition to asking the more traditional question of which group recalls more units, one can ask if groups of subjects tend to recall the same units independently of differences in the level of recall. A variety of hypotheses can be compared by correlating the values provided by each hypothesis for each unit of the story with the percentage recalled values of those units. Possible examples might include hypotheses based on rated importance, redundancy as measured by the CLOZE technique, serial position of the units (Rubin, 1978), or more complex process models (Kintsch & van Dijk, 1978).

Examining which units are recalled should be especially revealing for clinical populations. Differences in the total amount recalled can be attributed to differences in motivational level, attention, and other factors which often only incidentally relate to the cognitive processes under study. Differences in which units are recalled are less likely to be affected by such factors (Rubin, 1978). Moreover, knowing which units are recalled could provide a rich data base from which to draw inferences about processing. Specific hypotheses, such as the examples given above, or hypotheses based on theories about particular clinical groups, could be formulated to probe differences among groups. In this manner, differences in which units are most memorable could be predicted by most models of processing that specify a particular form of deficit such as a failure to appreciate the redundancies or gist of a story. If little difference is found in which units are recalled, one would argue that the processing of text is similar in the groups being compared and could be understood without reference to the particular clinical population being studied.

METHOD

Subjects

The 19 retarded children selected from schools around Durham, NC had IQs between 55 and 70, reading levels at or above the preprimer level, and chronological ages between 10 yr., 5 mo. 15 yr., 8 mo., with an average age of 13 yr., 0 mo. The 19 nonretarded children from the same school systems were

selected in an attempt to match mental ages. Their ages ranged from 6 yr., 6 mo. to 9 yr., 6 mo. with an average age of 8 yr., 1 mo. There were 13 males and 6 females in both groups. The sampling, which was done without regard to race, yielded 17 black and 2 white retarded children, and 6 black and 13 white nonretarded children.

Materials

Three stories were designed which would be age appropriate in style and content, similar in length, and which would require the subjects to fill in a variety of inferences to appreciate the story fully. The last criterion was included to maximize the chance of finding differences between the two groups. That is, we did not want stories that were such good figures they would compel all subjects to recall exactly the same thing.

One story was modified from Paris (1975) and two others were specially written for the study. The three stories, titled the Bozo, Airplane, and Linda stories, were divided into 44, 34, and 37 logical units, respectively, based on the scoring of pilot data. The stories divided into units are presented in Bacon (1978).

Procedure

The subjects, who were tested individually, were asked to listen to each story so that they could retell it in their own words immediately after it was read to them. A minimum level of motivation was maintained by verbal praise and by the previously announced reward of a penny, a piece of candy, or a piece of fruit following each recall. Tapes of each session were transcribed for scoring. The order of the stories was randomized for each subject.

RESULTS

All stories were scored for meaning, not exact wording. Reliability between two scorers for the number of units recalled from a sample of 33 protocols was .96. The percentages of units recalled on all three stories combined for the retarded and nonretarded groups are 35% versus 36%. The percentage of units recalled for the Bozo, Airplane, and Linda stories considered individually for the retarded and nonretarded group are 34% versus 34%, 32% versus 43%, and 39% versus 34%, respectively. A subject population (2) by story (3) analysis of variance yielded no effect of subject population ($F_{1,36} = 0.21$) or story ($F_{2,72} = 0.57$). Subjects approximately matched on mental age did not differ over-all on amount of recall as measured. The interaction of subject populations and story, however, was significant ($F_{2,72} = 3.88, p < .05$) because, as the means show, one story was recalled equally well by both groups, one story was recalled better by the nonretarded children, and one story was recalled better by the retarded children. Tukey and Newman-Keuls *post hoc*

tests were not significant for any of the differences between subject populations for the individual stories, though less conservative individual t tests were for one story, the Airplane story ($t_{36} = 2.27, p < .05$). The observation that one story might show a statistically significant difference when considered alone is consistent with work by Luftig and Johnson (1982).

Of more interest here than how many units were recalled by the two groups, however, is which units of the stories were recalled. The correlations between the percentage of retarded and nonretarded subjects recalling each of 44, 34, and 37 units of the Bozo, Airplane, and Linda stories are .832, .784, and .619, respectively. The reliabilities (Cronbach's alpha; Cronbach, 1951) for the percentage-of-subjects-recalling-each-unit values for the three stories are .804, .790, and .807 for the retarded and .836, .860, and .738 for the nonretarded sample. The correlations between the groups' approach the groups' reliabilities: averaging over stories, .745 versus .806. That is, the two groups correlate with each other nearly as highly as they would be expected to correlate with new samples of 19 subjects drawn from their own populations. This is not a negative finding due to a lack of sensitive measures. There is great variability among the units in their level of recall, variability that the reliability measures show is not random. Moreover, if the units were considered independent observations, the correlations between the groups would all be significant at the .001 level.

Another way to examine whether different groups recall the same items is Guttman scaling. This method examines individual subject's as opposed to group's behavior and allows testing for statistical significance, which correlations over nonindependent units do not (Rubin, 1978, 1980). Given the group's rank ordering of units from most to least likely to be recalled and the number of units individual subjects recall, we can predict which units will be recalled with an accuracy equivalent to the Guttman coefficient of reproducibility (Kenny & Rubin, 1977). By combining the data from the retarded and nonretarded groups we can ask whether the rank ordering of units from most to least likely to be recalled from the two individual groups are better predictors than the rank ordering from the groups combined (Rubin, 1978). All coefficients of reproducibility for the individual groups are greater than their chance levels, indicating that the recalls are Guttman-scalable, all $ps < .05$ using Green's (1956) method as the basis of calculating chance levels (Kenny & Rubin, 1977) and standard errors (Chilton, 1969). Of more importance, there is only a slight decrease in scalability when the retarded and nonretarded subjects are combined. In particular, the coefficients of reproducibility for the Bozo, Airplane, and Linda stories are .746, .749, and .772 for the retarded subjects, .794, .759, and .775 for the nonretarded subjects, and .767, .740, and .754 for both groups combined. That is, given the rank order of units from

most to least likely to be recalled based on all 38 subjects from both groups combined, we can predict which units individual subjects will recall nearly as accurately as we can from the rank order of the subject's own group: averaging over stories, 75.4% versus 76.6%.

DISCUSSION

Results of the comparisons of the two groups indicate clearly that retarded children are able to comprehend and remember the most salient units of a story the same as nonretarded children of a similar mental age. This finding is consistent with developmental studies (Brown & Smiley, 1977; Christie & Schumacher, 1975; Mandler & Johnson, 1977), studies of retarded children Lufitg & Johnson, 1982; Paris, *et al.*, 1974), and studies of other clinical populations (Graybeal, 1981; Rubin, *et al.*, 1981).

What does this similarity in which units are recalled mean? While retarded children and their controls may differ in how much they remember from some stories, the retarded and control children tend to recall the same units most often. Differences in total amount can be attributed to many variables not directly related to cognitive processes, such as motivational differences. Differences, or the lack of them, in the particular units that tend to get recalled, are harder to explain without recourse to how the individual units are processed. That is, if retarded children are recalling most often the same units as their controls, then they must be interacting with the stimulus in a fashion similar to their controls. While this work is consistent with work showing that retarded children are sensitive to three levels of rated importance or salience, it goes beyond this claim. Because the correlation between the two groups of subjects approaches their reliability, whatever variable can be shown to have an effect on recall in one group (e.g., imagery, grammatical structure, role in a story grammar, as well as rated importance or salience) will have a similar effect on recall in the other. That is, the unit analysis of discourse used here provides a tool for measuring verbal behavior, which imposes minimal bias on the structure of recall.

REFERENCES

- BACON, E. H. Visual and verbal cues in an oral comprehension task for retarded and nonretarded children. (Doctoral dissertation, Univer. of North Carolina in Chapel Hill, 1977) *Dissertation Abstracts International*, 1978, 39, 01A.
- BROWN, A. L., & SMILEY, S. S. Rating the importance of structural units of prose passages: a problem of metacognitive development. *Child Development*, 1977, 48, 1-8.
- CHILTON, R. J. A review and comparison of simple statistical tests for scalogram analysis. *American Sociological Review*, 1969, 34, 238-245.
- CHRISTIE, D. J., & SCHUMACHER, G. M. Developmental trends in the abstraction and recall of relevant versus irrelevant thematic information from connected verbal material. *Child Development*, 1975, 46, 598-602.
- CRONBACH, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, 16, 297-334.

- GRAYBEAL, C. M. Memory for stories in language-impaired children. *Applied Psycholinguistics*, 1981, 2, 269-283.
- GREEN, B. F. A method of scalogram analysis using summary statistics. *Psychometrika*, 1956, 21, 79-88.
- KENNY, D. A., & RUBIN, D. C. Estimating chance reproducibility in Guttman scaling. *Social Science Research*, 1977, 6, 188-196.
- KINTSCH, W., & VAN DIJK, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, 85, 373-394.
- LUFTIG, R. L., & JOHNSON, R. E. Identification and recall of structurally important units in prose by mentally retarded learners. *American Journal of Mental Deficiency*, 1982, 86, 495-502.
- MANDLER, J. M., & JOHNSON, N. S. Remembrance of things parsed: story structure and recall. *Cognitive Psychology*, 1977, 9, 111-151.
- PARIS, S. G. Integration and inference in children's comprehension and memory. In F. Restle, R. M. Shiffrin, H. J. Castellan, H. R. Lindman, & D. B. Pisoni (Eds.), *Cognitive theory*. Vol. 1. Hillsdale, NJ: Earlbaum, 1975. Pp. 223-246.
- PARIS, S. G., MAHONEY, G. J., & BUCKHALT, J. A. Facilitation of semantic integration in sentence memory of retarded children. *American Journal of Mental Deficiency*, 1974, 78, 714-720.
- RUBIN, D. C. A unit analysis of prose memory. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 599-620.
- RUBIN, D. C. 51 properties of 125 words: a unit analysis of verbal behavior. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 736-755.
- RUBIN, D. C., OLSON, E. H., RICHTER, M., & BUTTERS, N. Memory for prose in Korsakoff and schizophrenic populations. *International Journal of Neuroscience*, 1981, 13, 81-85.
- STEIN, N. L., & GLENN, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing*. Norwood, NJ: Ablex Publ. Corp., 1979. Pp. 53-120.
- STONE, C. P., GIRDNER, J., & ALBRECHT, R. An alternate form of the Wechsler Memory Scale. *Journal of Psychology*, 1946, 22, 199-206.

Accepted August 22, 1983.