Schizophrenic Recall and Contextual Constraint

Michael L. Raulin and Loren J. Chapman
University of Wisconsin—Madison

This study compared the effects of contextual constraint and list length on short-term recall of word lists by 24 chronic schizophrenics and 120 normal subjects. One subtest consisting of relatively short lists of low constraint and one subtest consisting of longer lists of high constraint were matched on mean, variance, and shape of the distribution of item difficulties, variance of subtest scores, shape of the distribution of subtest scores, and subtest reliability. These psychometrically matched subtests were used to compute a difference score of accuracy on low-constraint lists minus accuracy on high-constraint lists. On this difference score, schizophrenics scored lower than normal subjects with the same total accuracy scores. The direction of this difference is opposite to that found in four previous studies of the effects of contextual constraint on recall by schizophrenics. It was concluded that the findings of the previous studies are probably artifacts of the use of unmatched tasks and that schizophrenic deficit in recall is not increased more by an increase in contextual constraint than by a shortening of word lists. In fact, the data suggest that precisely the opposite may be true.

Miller and Selfridge (1950) defined contextual constraint as "the extent to which the choice of a particular word depends upon the words that precede it" (p. 177). They developed word lists at each of four different lengths at each of eight levels of constraint. They gave these 32 lists to college students and asked them to recall as many words as they could. The students recalled higher

percentages of words at the higher levels of constraint.

Several investigators have studied the effects of contextual constraint on schizophrenic recall. These studies have yielded conflicting findings. Lewinsohn and Elwood (1961), using the passages that were developed by Miller and Selfridge, found that chronic schizophrenic subjects showed a greater deficit, that is, a greater inferiority to normal subjects, in recall of the high-constraint lists than in recall of the low-constraint lists. These investigators inferred that the schizophrenics were less able than normal subjects to benefit from increased contextual constraint. Lawson, McGhie, and Chapman (1964) and Levy and Maxwell (1968) replicated these results. However, Raeburn and Tong (1968), using the same lists and the same procedure, were unable to replicate these results for either of two schizophrenic samples.

Other investigators have constructed high- and low-constraint tasks in other ways. Truscott (1970) used four types of word lists—normal, anomalous, semantically related, and random word strings. She compared schizophrenic and normal subjects on recall of all four and found the greatest performance difference between the two groups on the normal passages.

It is difficult to interpret any of these studies of schizophrenic response to contextual constraint because the tasks at the various levels of constraint were not matched on psychometric characteristics which affect their discriminating power. The discriminating power of a test refers here to the power of the test to yield a difference in mean score between the more able and the less able subjects. If two groups of subjects of different ability levels are given two tests of different discriminating power, the groups will differ more on the more discriminating test. Tests which have the same true-score variance have the same discriminating power in this sense of the term. True-score variance is the product of test reliability and obtained-score variance. Chapman and Chapman (1973a, 1973b) have advocated that tasks be matched on mean, variance, and shape of the distribution of item difficulties, shape of the distribution of test scores, and test reliability to assure equal true-score variance and, hence, equal discriminating power.

Lord (1952) has pointed out that for free-response items, the 50% level of difficulty yields the greatest reliability and variance and, hence, the greatest discriminating power. In evaluating studies which used unmatched tasks, one may often infer which task had the greater discriminating power from the tasks' difficulty levels. However, inferences about discriminating power cannot be made from difficulty level alone in studies of contextual constraint. Words in high-constraint lists tend to be remembered in clusters. This clustering effect raises both the variance and reliability in the high-constraint lists and, hence, raises their discriminating power. It follows that inferences about discriminating power of tasks in studies of contextual constraint should be made from reliability and variance rather than difficulty alone unless the task that is closest to the middle range of difficulty is also the task with the greatest contextual constraint.

Truscott (1970) gave sufficient information that the reader may make reasonable guesses as to the direction of the probable artifact that resulted from her failure to match tasks. Her normal word lists were much closer than the other types of word lists to the 50% level of difficulty for the two groups combined. The normal word lists would be expected to have the greatest discriminating power both because of their difficulty level and their higher contextual constraint. Consequently, the divergence of groups with increasing contextual constraint found by Truscott may well be attributable to the use of unmatched tasks and not to an inability of schizophrenics to benefit from increased contextual constraint. From the published data of Lewinsohn and Elwood (1961), Lawson et al. (1964), Levy and Maxwell (1968), and Raeburn and Tong (1968) one cannot infer as readily which task had the greatest discriminating power.

The present study was designed to determine the relation between level of contextual constraint and adequacy of recall by schizophrenic and normal subjects, using tasks that were matched on discriminating power for normal subjects. Two levels of contextual constraint were used: Level 1 (low constraint) and Level 5 (high constraint), as defined by Miller and Selfridge (1950). Levels 1 and 5 were chosen because previous investigators found the widest disparity in differential performance of the schizophrenic and normal subjects at these levels (Lawson et al., 1964; Levy & Maxwell, 1968; Lewinsohn & Elwood, 1961).

Since high-constraint lists are inherently less difficult than low-constraint lists, they can be matched on difficulty only by introducing a second variable. The variable chosen was length of list. The high-constraint lists were made longer than the low-constraint lists so that they could be matched on mean percentage accuracy for normal subjects. Length of list was chosen as the control variable because one would not think that it taps an important schizophrenic deficit. With this design change, the hypothesis becomes one of differential deficit. The hypothesis is that schizophrenic deficit in recall is increased more by an increase in contextual constraint than by a shortening of word lists. This hypothesis follows from the expectation that schizophrenics will be less able to profit from increased contextual constraint than from a shorten-

ing of the word lists on tasks matched so that normal subjects profit equally from both variables.

METHOD

Subjects

The normal subjects included 120 males (94 prison inmates and 26 firemen) ranging in age from 18 to 49. Prison inmates were used in order to obtain subjects whose accuracy scores would be low. Inmates were screened for intelligence using the Wide Range Intelligence Scale. Only inmates with IQ scores of 100 and below were chosen. No subject was included in this group if he had a history of any psychotic disturbance or psychiatric hospitalization. This sample contained 77 white subjects and 43 subjects who were either black, Mexican American, or American Indian. At the time of testing, 64 subjects had never been married, 36 were currently married, 14 were either divorced or separated, and no information was available for 6 subjects. The mean age and education for these subjects were 26.6 and 10.5 years, respectively.

Twenty-four chronic schizophrenics, 12 males and 12 females, ranging in age from 23 to 62, were tested. No patient had received electroshock therapy in the 3 months prior to testing, and all but one of the patients had been off antipsychotic medication for 4 weeks or more. (This one patient had been withdrawn from drugs 5 days prior to testing.) The patients were drug free because they were serving as control subjects in biomedical research. This sample of schizophrenics consisted of 22 white and 2 black subjects. At the time of testing, 19 patients had never been married, 3 were married, and 2 were divorced. The mean age, education, and length of hospitalization were 41.8, 11.3, and 16.0 years, respectively.

Procedure

Preliminary development of the lists. Using the method developed by Miller and Selfridge (1950), 60 lists were constructed with the aid of 50 college student volunteers. The students were tested sequentially, one at a time. High-constraint lists were created by giving each student sequences of words, each of which was four words long. The examiner presented one sequence of words at a time and urged the student to "build a sentence around them." The starter sequences for the first subject were drawn from current fiction. The word immediately following the four-word sequence in the student's sentence was added to the growing list. The last four words of the list became the sequence for the next student. The high-constraint lists for the study were portions of these lists. The low-constraint lists were created by placing all the words from a number of other high-constraint lists into a pool and drawing out words one at a time, without replacement and without restrictions. The high-constraint lists ranged in length from 8 to 16 words, while the low-constraint lists ranged in length from 6 to 10 words.

The word lists were recorded on magnetic tape. A male speaker read the words at a rate of 40 words per minute at a uniform volume and in a near monotone voice. A short tone signaled the end of each list.

Experimental procedure. Each experimental subject was asked to listen to the lists and then, after the tone, to repeat as many of the words as he or she could remember. The subject's responses were recorded on magnetic tape for later scoring. All subjects received the lists in the same order.

The experimental instrument intermixed 10 high-constraint and 10 low-constraint lists which were selected on the basis of appropriate pilot work. These lists were preceded by recorded instructions and two practice lists. A 5-minute break was taken half way through the test.

The data from the normal sample of 120 subjects were scored for total accuracy and divided into two samples of 75 and 45 subjects each on the basis of these accuracy scores. Total accuracy scores were computed by simply adding up the total number of words correctly identified by each subject across all of the lists. These two samples were chosen so that they had virtually identical distributions of total accuracy scores. The 75 subjects became the standardization sample and the remaining 45 subjects became the cross-validation sample. The high-constraint and low-constraint tasks were matched on the psychometric characteristics affecting discriminating power using the data from the standardization sample. The cross-validation sample was used both to check the matching of the two tasks and as a normal comparison group for the analysis of the schizophrenics' data.

Interscorer reliability. Interscorer reliability for accuracy was very high. The product-moment correlations based on 28 inmates' responses, scored independently by two raters were .98 for the low-constraint subtest and .99 for the high-constraint subtest.

Matching the subtests. For matching the subtests, each list was treated as an item. A high-constraint and a low-constraint subtest were chosen so that they were matched as closely as possible on reliability, mean, variance, and shape of the distribution of item difficulties, and shape of the distribution of test scores for the normal standardization sample. (Item difficulty is the mean accuracy score for an item.) To achieve the match on test reliability and variance despite the tendency of high constraint words to cluster in recall, more lists were used for the low-constraint subtest. The matched subtests included 6 of the 10 low-constraint items and 5 of the 10 high-constraint items. Table 1 presents the psychometric characteristics of these matched subtests for both the normal standardization sample and the cross-validation sample. The estimated true-score variances for the normal standardization sample were fairly close, .51 for the low-constraint subtest and

TABLE 1

PSYCHOMETRIC PROPERTIES OF TWO MATCHED TESTS

Psychometric characteristic	Group			
	Standard-ization (n = 75)	Cross-valida-tion (n = 45)	Schizo-phrenic (n = 24)	Low-scoring normal (n = 25)
Mean accuracy score				
Low	3.72	3.54	2.64	2.82
High	3.15	3.15	2.60	2.35
Mean item difficulty[a]				
Low	.62	.59	.44	.47
High	.63	.63	.52	.47
Variance of item difficulty				
Low	.013	.014	.008	.009
High	.012	.010	.016	.010
Subtest variance				
Low	.65	.54	1.15	.49
High	.53	.56	.60	.45
Coefficient alpha[b]				
Low	.79	.77	.85	.69
High	.80	.85	.66	.73
Estimated true score variance				
Low	.51	.42	1.28	.34
High	.43	.48	.40	.33

Note. Low and high refer to level of contextual constraint.
[a] Mean item difficulty is the average proportion of words per list correctly identified.
[b] Coefficient alpha was used to estimate reliability.

.43 for the high-constraint subtest. These values were a bit closer for the cross-validation sample. The matching on shape of the distribution of item difficulty appeared somewhat less satisfactory.

RESULTS

Schizophrenic Performance

The schizophrenics, as expected, scored lower than the normal subjects on both subtests. Table 1 shows their performance data. The schizophrenics appeared to approach more closely the performance of the cross-validation subjects at the higher levels of contextual constraint than at the lower levels.

Before computing significance tests, small deviations in mean and variance between the two subtests were corrected by converting all raw scores to standard scores on the basis of the mean and variance for the standardization sample. In order to test the study's hypothesis, a *t* test was used to compare the schizophrenic and normal cross-validation groups on

the difference score of low constraint minus high constraint and was found to be nonsignificant, $t(67) = 1.38$, $p > .05$. This nonsignificant difference in performance between these two groups is in a direction that is opposite to that reported previously (Lawson et al., 1964; Levy & Maxwell, 1968; Lewinsohn & Elwood, 1961; Truscott, 1970).

The results obtained from this statistical analysis must be regarded as only approximate because of imprecise matching of the subtests. Tests that are closely matched on reliability and variance but not on the shape of the distribution of item difficulty for subjects at one level of accuracy may become unmatched on reliability and variance for subjects at a different accuracy level. This is likely to occur because differing numbers of items will move into the middle range of difficulty for low-scoring subjects, and the distribution of item difficulty affects reliability and variance. If the two tasks could have been matched precisely on the distribution of item difficulty, then the comparison of the schizophrenics with the cross-validation sample would have been an adequate test of the hypothesis. Since such matching was not possible, the possibility of a psychometric artifact remains. Therefore, a sample of 25 low-scoring normal subjects was drawn from the 120 normal subjects so that they were matched with the schizophrenic sample on overall accuracy.

The use of a sample of low-scoring normal subjects as a comparison group is an alternative to the approach of matching tasks. Tasks should be matched lest generalized poor performance yield a differential performance deficit on the two tasks. If the schizophrenic and normal samples are at the same overall accuracy level, this artifact could not occur. As seen in Table 1, the subtests appeared much more closely matched on true-score variance for this low-scoring normal sample than for the other normal groups. This apparent close matching for low-scoring subjects cannot be given full evidence because the sample size is much too small to give stable estimates of reliability and variance. However, as stated above, the logic of this analysis does not assume matched tasks. On the difference

score between the two subtests, the schizophrenic subjects differed from the low-scoring normal subjects, $t(47) = 2.33, p < .05$. Again the direction of this difference was opposite to that found by Lewinsohn and Elwood (1961), Lawson et al. (1964), Levy and Maxwell (1968), and Truscott (1970). The schizophrenic deficit in recall was increased less by an increase in contextual constraint than by a shortening of the word lists.

Effects of Age and Sex

Because the schizophrenic and normal samples differed on mean age and sex, we examined the data to determine if age or sex had any differential effect on performance. No differential effect on performance for age was found within the normal sample $(r = .03)$. Within the schizophrenic sample, sex made no difference in either the total performance, $t(22) = .04, p > .05$, or in a score of the difference between the two subtests, $t(22) = .25$, $p > .05$. Since no effect of age or sex was noted, it must be concluded that the differences between groups were due to the illness itself.

DISCUSSION

Clearly, these results cast doubt on the validity of the conclusions of a number of previous investigators that an important feature of schizophrenic deficit is a loss of ability to profit from increased levels of contextual constraint. We have already commented on the role of unmatched tasks as a potential artifact in those studies.

Another difference between the present study and the previous studies on the same topic is that none of the subjects in the present study were taking antipsychotic medication at the time of testing. Since the effects of these drugs are often as dramatic as the effects of the illness itself, and just as unpredictable, the results of the earlier studies could have reflected the effects of medication as well as the effects of schizophrenia.

The simple hypothesis that schizophrenics cannot profit from increased contextual constraint is inherently untestable because difficulty varies with constraint. Like most studies of differential deficit, our study used only a single control variable, that is, length of list. We chose a control variable which seemed unlikely to tap an important schizophrenic deficit. We recast the hypothesis in testable form to read that schizophrenic deficit in recall is increased more by an increase in contextual constraint than by a shortening of word lists. The original hypothesis was not confirmed and, in fact, the opposite differential deficit was found when the schizophrenics were compared with normal subjects of the same ability level. The results show that schizophrenic deficit in recall is not increased more by an increase in contextual constraint than by a shortening of the word lists. In fact, the results suggest that schizophrenic deficit in recall may actually be increased less by an increase in contextual constraint than by a shortening of word lists.

REFERENCES

Chapman, L. J., & Chapman, J. P. *Disordered thought in schizophrenia.* New York: Appleton-Century-Crofts, 1973. (a)

Chapman, L. J., & Chapman, J. P. Problems in the measurement of cognitive deficit. *Psychological Bulletin,* 1973, *79,* 380–385. (b)

Lawson, J. S., McGhie, A., & Chapman, J. Perception of speech in schizophrenia. *British Journal of Psychiatry,* 1964, *110,* 375–380.

Levy, R., & Maxwell, A. E. The effect of verbal context on the recall of schizophrenics and other psychiatric patients. *British Journal of Psychiatry,* 1968, *114,* 311–316.

Lewinsohn, P. M., & Elwood, D. L. The role of contextual constraint in the learning of language samples in schizophrenia. *Journal of Nervous and Mental Disease,* 1961, *133,* 79–81.

Lord, F. M. A theory of test scores. *Psychometric Monographs,* 1952, No. 7.

Miller, G. A., & Selfridge, J. A. Verbal context and the recall of meaningful material. *American Journal of Psychology,* 1950, *63,* 176–185.

Raeburn, J. M., & Tong, J. E. Experiments on contextual constraint in schizophrenia. *British Journal of Psychiatry,* 1968, *114,* 43–52.

Truscott, I. P. Contextual constraint and schizophrenic language. *Journal of Consulting and Clinical Psychology,* 1970, *35,* 189–194.