韓國心理學會誌
Korean Journal of Psychology
1987. Vol. 6, No. 1. 21－33

Comparison of Event-Frequency Judgment and Free Recall: Effects of Subject Age, Subject Gender, and Test Delay*

Yung Che Kim
Keimyung University
Daegu, Korea

Melvin H. Marx
Georgia State University
Atlanta, GA. USA

The experiment examined the covariation of the event-frequency judgment and free recall response as a function of subject's school grade, gender, and test delay. Subjects were shown a series of animal names, one on each slide, and were then asked both to recall them and to estimate the frequency occurred (0, 1, 3, or 5 times). Half of the subjects had the free-recall test first, the other half the frequency-judgment test first. The varibles of school grade (6th, 10th, and college students) and test delay (0, 1, or 7 days) produced similar results for free recall and frequency judgment; in general, the poorest performance was by the 6th grade students, the best performance was by the 10th grade students, and, as test delay being extended, performance was declined. The increment in frequency-judgment response from 6th graders to older subjects was the most significant result, presumably not reported previously. A variety of other effects and interactions were also found. The theoretical significance of the results were discussed with particular reference to its implications for the hypothesis that frequency judgment response would be automatic cognitive process.

People are surprisingly good at keeping track of the number of occurrence of different kinds of events. It seems to be true whether the event is the happening of a certain behavior (How many times have you been punished by your parents so far?) or the presentation of certain objects. Actually, a large body of empirical evidences support that people are sensitive to the frequency of events, including such as the presentation of a specific word (Hinzman, 1969), individual syllables occurring in a list of words (Underwood & Zimmerman, 1973), single letters (Attneave, 1953), the generation or a certain behavior (Williams & Durso, 1986), the reference to a superordinate concept (Barsalou & Ross, 1986), an even others (Williams & Durso, 1986). These results have been a major factor in the contention that event-frequency judgment or the encoding of frequency information is effortless, inflexible, and in-

* This research, which was a part of the comprehensive cooperative research project of two authors, was supported by a grant from the Korea Research Foundation, awarded to the first author. Requests for reprints should be sent to Yung Che KIM, Department of Psychology, Keimyung University, Daegu, the republic of Korea. Thanks are extended to the unidentified two reviewers.

accessible, in short, that is automatic, rather than an effortful cognitive function (Hasher & Zacks, 1979, 1984). Hasher and Zacks and others have presented numerous evidences relating to a number of criteria that they proposed as tests for automaticity of frequency encoding. To put briefly, the criteria are (a) sensitivity to the frequency of events shows little change with age and is emonistrable even in young children, (b) frequency encoding does not benefit from practice in tasks that require such encoding, (c) the accuracy of frequency of occurrence reports is not affected by the individual differences, (d) the variations in effortful encoding strategies or instructional strategies such as intention to code frequency does not affect the efficiency of frequency information registration, and (e), finally, encoding of frequency information is not affected by competing task loads or demands.

Evidence as to whether frequency information meets the criteria listed above is somewhat mixed. However, the criteria of no developmental trends has received relatively a large amount of empirical study. Hasher and Zacks have stated that "beyond the age of four or five, sensitivity to frequency may not change"(1984, p.1377). But, the evidence supporting the point is rather skimpy. A small number of studies have shown little difference in frequency judgment performance over the primary grades (Johnson, Raye, Hasher, & Chromiak, 1979; Hasher & Zacks, 1979, Experiment 1; Goldstein, Hasher, & Stein, 1983), and one study revealed minimal changes over grades 2, 4, and 6 and college students (Hasher & Chromiak, 1977, Experiment 1). There are some results that do show increments in frequency-judgment accuracy from kindergarten into the primary grades (Ghatala & Levin, 1973; footnote 3) point to methodological problems in these experiments. A relatively high degree of invariance in the frequency-judgment task has also been reported across the adult years, with college students usually compared with elderly subjects (Attig & Hasher, 1980; Hasher & Zacks, 1979, Experiment 2; Kausler & Puckett, 1980). In more recent reports, however, some "modest age deficit" has been found (Kausler, Lichty, & Freund, 1985; cf. also, Kausler, Lichty, & Hakami, 1984). There is no obvious way to resolve the discrepancies in these studies. Nevertheless, it is noted that spontaneous, extensive semantic processing will be performed on to-be-remembered information more by young adults than by children or elderly adults, because age has large effects on the type of processing that subjects perform.

This study was conducted to compare event-frequency judgment and free recall as a function of age and gender of subjects and delay of test. There does not appear to be any reported data that bridge the gap between the primary grades and the adolescent or early adult years, such as are provided in the present study. Moreover, there is a marked discrepancy between the well-documented increment with age in most cognitive functions, notably free recall, and the failure of event-frequency judgment to show any of such developmental trends. With respect to the test-delay variable, it was felt that additional data, beyond those early reported by Underwood, Zimmerman, and Freund(1971), would be desirable, particularly any showing interactions with age or gender.

METHOD

Experimental Design

The basic design was a 4 (age of subjects:6th graders, 10th graders, college students, adults) by 2 (gender of subjects) by 4 (test delay:0 day, 1 day, 7 days, 30 days) by 2 (test order:free recall or frequency judgment test first) factorial between subjects experiment with word frequency (0, 1, 3, 5 occurrences) varying within subjects. The test-order variable was included in order to give half of the subjects the free-recall test before they were given the target words in the frequency-judgment test so that an uncontaminated measure of recall could be obtained.

Subjects

A total of 1,152 subjects provided the data. The 12-year-old 6th grade subjects were obtained from two classes of sixth graders at each of three elementary schools and the 16-year-old subjects from two 10th grade (1st grade of the senior high school) classes at three senior high schools, all in the city of Daegu, Korea. The students were tested in their regular classroom settings. Classes were assigned randomly to the various treatment conditions. Males and females were mixed in the elementary school and college classes but were in separate classes in the senior high schools. The college students were sophomores from introductory psychology classes at the Keimyung University, Daegu, Korea. The classes were randomly assigned to the various treatment conditions. The adult subjects who were in their forties were selected from the parent pool of the 6th and 10th grade subjects.

After all of the data were collected and checked for completeness, a total of 179 subjects were then randomly discarded from 27 treatment conditions to provide an equal cell frequency at n=18, before the data were analyzed.

Word List

Because no word-frequency ratings like the Thorndike-Lorge (1944) was available for Korean words, an indepencent sample of 74 colege students was used to provide familiarity ratings for 55 animal names on a 5-point scale. Animal names were used rather than the more typical unrelated words in an effect to maintain the interest of the subjects in the experiment. On the basis of the survey, six sets of four animal names with approximately equal familiarity ratings were prepared. These sets consited of Korean-script words for frog, cat, fowl, and sparrow; dog, cattle, horse, and pig; pigeon, squirrel, magpie, and goat; elephant, crow, bear, and tiger; deer, seagull, lobster, and pheasant; and wolf, whale, fox, and giraffe. The four names in each set were randomly assigned to the four frequency categories (0, 1, 3, 5), so that each category consisted of six words. The resulting 54 words were randomly ordered, with the restriction that no word could follow itself. The mean number of intervening words for the items presented three times (F3) was 4.42 and the corresponding mean for the F5 words was 2.67. Two primacy buffers (ant, rabbit) and two recency buffer items (stock, hippopotamus) were also used, making a total of 58 words in the list.

Procedure,

The subjects were informed that they would be shown a series of animal names, one word at a time by slide projector, and were shown some sample slides (lizard, dee, butterfly). They were told that their task was to remember the words, and that some words would be shown more than once.

In the study phase, the list of 58 words were shown, one word at a time. Each word mounted on the slide was presented for 2 sec with a 2-sec inter-slide interval. Following the word presentation, indi-vidudual test sheets were distributed and the subjects were cautioned to wait quietly for the next phase of the experiment.

The two tests were then administered in one of the two orders (free-recall or frequency-judgment test first.) The free-recall test was introduced with the usual instructions to write down as many of the names as could be remembered, without regard for the order in which they had appeared. For the frequency-judgment test the subjects were told that the test sheet contained the names of some of the animals that had been shown along with some that had not been. They were also informed that the maximum number of occurrences had been 5. They were instructed to write down the number of times that they thought each of the animal names had been shown. About two minutes were allowed for each test. The administration time interval between the two tests were about 5 minutes.

RESULTS

Correct Responses

The number of correct responses in each of the free-recall and frequency-judgment test was calculated in order that a direct comparison could be made between free-recall (FR) and frequency-judgment (FJ) performance.

The frequency-judgment data analysed and presented in Table 1, 3, 5, and 6 are all based on the 576 subjects who had their frequency-judgment test first. Similarly, only the data for the other 576 subjects who had their free-recall test first, so that it could not be inflated by the additional study provided by the frequency-judgment test, are included in Table 2 and in subsequent free-recall analyses. Table 1 and 2 shows the major results. Mean numbers of correct frequency judgments for the four age groups and the four test delays for all subjects are shown in Table 1 and mean number of correct word recall for the four age groups and the four test delays are presented in Table 2. With the criterion of the number correct FJs, the main effects of age and delay were all statistically reliable with $F(3, 544)=8.27$, $MSe=5.21$, and $F(3, 544))=72.46$, $MSe=5.21$, $p<.01$, respectively. Furthermore, the interaction effect between age and test delay was also reliable, $F(9, 544)=4.12$, $MSe=5.21$, $p<.01$, with a markedly greater decline in number correct FJs occurring sooner for the 6th graders, as suggested in Table 1. When the Scheffe's post hoc test was applied, both the overall increment in mean correct FJ scores from the 6th grade group to the 10th grade group (9.61-8.32=1.29) and its mean decrement from the 10th grade group to the combined group of the college students and adults group (9.61-[8.96+8.65]/2=0.805) were all reliable with $F(3, 572)=15.74$, $p<.01$, and $F(3, 572)=12.11$, $p<.01$, respectively. But, no other group differences were found reliable.

The counterpart scores for the FR measure are presented in Table 2. For FR responses again, both age and delay main effects were reliable, age $F(3, 544)=26.63$, $MSe=6.25$, $p<.01$, and test delay $F(3, 544)=41.01$, $MSe=6.25$, $p<.01$, test delay $F(3, 544)=41.01$, $MSe=6.25$, $p<.01$. Also the age and test delay interaction effect was reliable, $F(9, 544)=29.14$, $p<.01$; but the sex main effect was not reliable, $F(1, 544)=1.39$, $p>.01$. The Scheffe's test was also applied. The mean increment in FR scores from the 6th grade to the 10th grade group (13.32-11.37=1.95) was reliable, $F(3, 572)=33.76$, $p<.01$, as was the mean decrement from the 10th grade group to the col-

Table 1 Mean numbers of correct frequency judgments for the four age groups and the four test delays (S.D.s in parentheses).

Delay (days)	Age: 6th graders	10th graders	College students	Adults	Total
0	10.78(2.09)	11.64(2.38)	10.00(2.17)	11.20(2.93)	10.91(2.39)
1	7.56(2.21)	10.25(2.61)	9.83(2.05)	9.47(2.07)	9.28(2.24)
7	8.61(2.70)	8.33(2.66)	8.75(2.17)	7.39(2.36)	8.27(2.48)
30	6.34(2.03)	8.20(2.35)	7.25(2.55)	6.53(1.72)	7.08(2.16)
Sub-total					
Male	8.17(2.28)	9.08(2.37)	8.63(2.09)	8.06(2.27)	8.49(2.26)
Female	8.47(2.23)	10.13(2.63)	9.29(2.37)	9.24(2.26)	9.28(2.38)
Total	8.32(2.26)	9.61(2.50)	8.96(2.23)	8.65(2.27)	8.89(2.32)

Table 2 Mean numbers of correct word recalls for the four age groups and the four test delays (S.D.s in parentheses).

Delay (days)	6th graders	10th graders	Age College students	Adults	Total
0	12.03(1.69)	14.75(2.04)	12.75(1.82)	11.69(2.44)	12.81(2.00)
1	12.44(2.54)	13.97(2.94)	13.50(2.02)	11.23(2.75)	12.79(2.56)
7	11.36(2.64)	12.83(2.90)	12.09(2.42)	9.48(2.68)	11.44(2.66)
30	9.67(2.35)	11.75(2.33)	10.03(2.67)	8.67(2.10)	10.03(2.36)
Sub-total					
Male	11.78(2.34)	12.54(2.51)	11.77(2.30)	9.90(2.38)	11.50(2.40)
Female	12.02(2.26)	13.07(2.59)	12.42(2.16)	10.63(2.60)	12.04(2.39)
Total	11.37(2.30)	13.32(2.55)	12.10(2.32)	10.27(2.49)	11.77(2.40)

lege students group, $F(3, 572)=13.21$, $p<.01$. But, neither the other individual group differences nor the difference between the 6th grade group and the combined group of the college students and adults reached the value to be reliable at the significance level of $p=.01$. While the group difference between the 10th grade group and the college students group was reliable in FR, it was not in FJ measure: It was the only discrepancy between these two measures of worth to be noted.

Figure 1-A provides a direct comparison of the mean numbers of correct FJ and FR responses for the age variable. Presented in this way, FJ and FR mean scores show a remarkably similar variation with age. In each case, the highest performance level occurred in the 10th grade subjects. The decremental change of FR compared to the FJ scores from the college students group to the adult group was somewhat drastic.

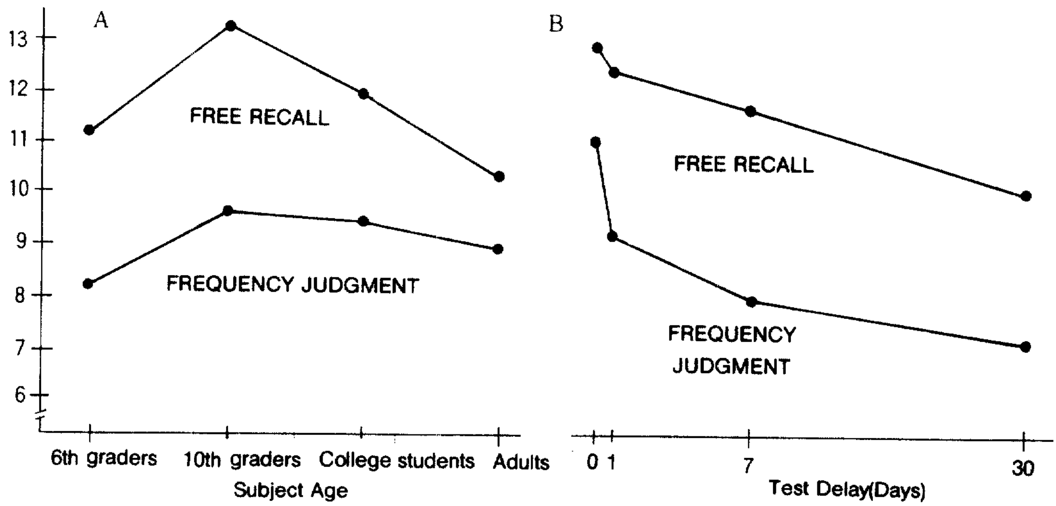

Figure 1 Mean numbers of correct free recall and frequency judgment responses as a function of (A) subject age and (B) test delay.

The comparable graph for the test-delay variable is shown in Figure 1-B. Here again there was some similarity of function between the two measures. It was evident that both the FJ and FR scores fell off very much similarly with increasing test delay. However, the fact that the FJ scores fell off more rapidly with 1-day delay than did the FR scores could be noted. The respective Scheffe' test run on the FJ measure showed that the decrement from 0-to 1-day delay (10.91-9.28=1.63), $F(3, 572)=32.81$, $p<.01$, that from 1-day to 7-day (9.28-8.27=1.01), $F(3, 572)=12.59$, $p<.01$, and the from 7-day to 1-month test delay (8.27-7.08=1.19), $F(3, 572)=17.40$, $p<.01$, were all reliable. The Scheffe' test for the FR measure also revealed the reliable decrement from 1-to 7-day test delay mean scores (12.79-11.44=1.35), $F(3, 572)=16.80$, $p<.01$, from 7-day to 1-month test delay mean scores (11.44-10.03=1.41), $F(3, 572)=10.76$, $p<.01$. However, the mean FR change from the immediate to the 1-day delayed test was not reliable, $F<1.00$.

Gender

In the FJ data, there was a reliable gender main effect, $F(1, 544)=17.60$, $MSe=5.21$, $p<.01$;but none of the interaction effects with gender variable were statistically reliable. As shown in Table 1, female subjects were consistently superior to male subjects, with mean score of 9.28 compared to 8.49 of the counterpart group.

Here also, suprisingly similar results were obtained for the FR data. The correct recalls of female group were consistently superior to male scores, but their difference was not large enough to be reliable, $F(1, 544)=1.39$, $MSe=6.25$, $p>.01$, with the mean score of 12.04 for the female and 11.50 for the male group. Also, none of the interaction effects with the gender variable were reliable. It should specially be mentioned that there was no interaction effect of gender with age for either FJ or FR measure, $F(3, 544)=1.07$ for FJ and $F(3, 544)=1.92$ for FR measure.

Item Frequency

Lack of the control for specificity of items in this study has not been a factor in any of the analyses thus far described, because item frequency has not been involved. But, in this section, some interesting apparent effects of the item-frequency variable are presented with some caution that there could be some confounding effects of specific items and frequencies.

The mean numbers of correct frequency judgments as a function of subject age and item frequency are

Table 3 Mean numbers of correct frequency judgments as a function of subject age and item frequency (S.D.s in parentheses).

Item Frequency	Age 6th graders	10th graders	College students	Adults	Total
0	4.25(1.20)	4.75(1.27)	4.58(1.31)	4.20(1.27)	4.45(1.26)
1	1.80(1.30)	2.45(1.43)	1.91(1.45)	1.75(1.36)	1.98(1.39)
3	1.82(1.15)	1.76(1.21)	1.90(1.12)	1.85(1.38)	1.83(1.22)
5	0.48(0.81)	0.65(0.91)	0.59(0.93)	0.86(1.02)	0.64(0.92)
Sub-total					
Male	2.04(2.28)	2.27(2.37)	2.16(2.09)	2.02(2.27)	2.12(2.26)
Female	2.12(2.23)	2.53(2.63)	2.32(2.37)	2.31(2.26)	2.32(2.38)
Total	2.08(2.26)	2.40(2.50)	2.24(2.23)	2.16(2.27)	2.22(2.32)

Table 4 Mean numbers of correct recalls as a function of subject age and item frequency (S.D.s in parentheses).

Item Frequency	Age 6th graders	10th graders	College students	Adults	Total
1	3.51(1.10)	3.65(1.22)	3.27(1.24)	2.95(1.41)	3.35(1.24)
3	3.91(1.27)	4.49(1.11)	4.06(1.13)	3.39(1.17)	3.96(1.17)
5	4.48(1.04)	4.67(1.08)	4.77(1.08)	3.94(1.01)	4.46(1.05)
Sub-total					
Male	3.93(2.34)	4.18(2.51)	3.92(2.30)	3.30(2.38)	3.83(2.40)
Female	4.00(2.26)	4.35(2.59)	4.14(2.16)	3.54(2.60)	4.01(2.39)
Total	3.97(2.30)	4.27(2.55)	4.03(2.23)	3.42(2.49)0	3.92(2.40)

Table 5 Mean numbers of correct frequency judgments as a function of test delay and item frequency (S.D.s in parentheses).

Delay (days)	Item Frequency 0	1	3	5	Total
0	5.48(0.89)	3.14(1.61)	1.66(1.19)	0.63(0.85)	10.92(2.39)
1	4.68(1.29)	2.23(1.58)	1.76(1.16)	0.62(0.96)	9.28(2.24)
7	4.02(1.45)	1.53(1.31)	1.99(1.25)	0.74(1.06)	8.27(2.48)
30	3.59(1.42)	1.01(1.07)	1.91(1.26)	0.58(0.81)	7.08(2.16)
Sub-total					
Male	4.30(1.26)	1.74(1.30)	1.88(1.22)	0.57(0.93)	8.49(2.26)
Female	4.59(1.26)	2.21(1.48)	1.78(1.21)	0.71(0.90)	9.28(2.38)
Total	4.45(1.26)	1.98(1.39)	1.23(1.22)	0.64(0.92)	8.89(2.32)

presented in Table 3. The significant interaction effect of age and item frequency, $F(6. 816)=11.49$, $MSe=1.74$, $p<.01$, was primarily due to the relatively wide dispersion of the FJ scores at the frequency-one (F1) level compared to the other levels.

The comparable free recall data are presented in Table 4. Again, the reliable interaction effect of age with item frequency variable was found with $F(6, 816)=12.01$, $MSe=1.41$, $p<.01$. The primary factor of this interaction effect was the marked superioritty in free recall of the higher-frequency items by the adult and, particularly, the college students group.

For the FJ scores, the only reliable ANOVA result of interest was the interaction effect of item frequency and test delay, $F(6, 816)=27.76$, $MSe=2.10$, $p<.01$. The interaction effect again seems to be related to the relatively invariant changes of F3 and F5 over the test delay period compared with the F0 and F1 frequency items.

Correlational Analysis of Frequency Judgments

The correct-response measure utilized only a small portion of the responses, especially for the higher-frequency items, as shown in Table 3. So, a more comprehensive statistical analyses were applied to the event-frequency judgments. Pearson product-moment correlational coefficients were computed between real and estimated frequencies for each subject. Correlation coefficient would assess the discriminability of the judgments rather than some absolute deviation that the judgments may have had from the actual frequencies. These rs were then converted into Fisher z scores to permit arithmetic operations and ANOVAs. This method of analysis was expected to be useful in confirmation of the frequency results which were based on correct responses alone, because these analyses presented in the previous sections were practically concerned with the deviation of the judged from actual frequencies. The product-moment correlation coefficients between the frequency judgments and actual frequencies for the four age groups and four test-delays are shown in Table 6.

The age main effect was reliable, $F(3, 544)=10.02$, $MSe=0.59$, $p<.01$. Once again, the variation of rs with age was found remarkably similar to that of FR mean scores which was shown in Figure 1. This suggests that FR mean scores and the discriminability of the frequency judgments show very similar variation pattern over subject ages. The mean correlation coefficients was .580, .655, .640, and .585, for the group of 6th grade, 10th grade, college students, and adults respectively. The test-delay variable main effect was also reliable, $F(3, 544)=77.02$, $MSe=0.05$, $p<.01$. As can be expected, the mean correlation coefficients were lowered as the test-delay period extended. Regardless of the experimental conditions, subjects made more accurate judgments on the no-delay frequency-judgment test, as can easily be expected. The interaction effect of age and test-delay variable was also reliable, $F(9, 544)=3.54$, $MSe=0.52$, $p<.01$. The most interesting facet of the frequency judgments by 6th graders was their marked inferiority after the 1-day delay. Gender was an ineffective variable in the analyses, $F<1.00$, nor did it interact with the age or the test-delay variable.

DISCUSSION

The remarkable similarity of the relationship of frequency judgment and free recall to the age variable is the most important result found in this study. These results which used the correct-response measure of FJ and FR were further confirmed by the correlational analysis. The increment in performance on frequency estimation matching the commonly observed increment in free recall as a function of age has not been reported previously. As already noted in the intorduction part, Hasher and Zacks (1984) extrapolate the null results of frequency judgment studies using mainly younger subjects (with college students included in one case), in the absence of any single study in which subjects covering the age span we have used were

Table 6 Product-moment correlations between frequency judgments and actual frequencies for the four age groups and four test delays.

Delay (days)	6th graders	10th graders	Age College students	Adults	Total
0	0.715	0.768	0.723	0.713	0.730
1	0.513	0.650	0.693	0.643	0.630
7	0.613	0.613	0.600	0.503	0.585
30	0.443	0.545	0.513	0.435	0.485
Sub-total					
Male	0.595	0.625	0.640	0.575	0.610
Female	0.565	0.675	0.640	0.595	0.625
Total	0.580	0.655	0.640	0.585	0.615

tested. The finding of change with ages by both measures of the correct-response and the correlational coefficient suggests the need to make some qualification in the proposition that there are no essential differences in the ability to make event-frequency judgments over the life span. Confidence in these results is strengthened by their contrast with the complete ineffectiveness of the same variables, notably age and test delay, when the numerical values of the frequency judgments were analyzed.

Regarding the Hasher and Zacks(1979, 1984) hypothesis that frequency judgment is an automatic cognitive function, what implications do these results have? Not very much in the way of definitive implications, in our opinion. Only if one specifically includes an assumption of uniformity of frequency judgment ability over the life span as a part of the automaticity proposition would these results be directly relevant. It seems very reasonable to assume that automatic cognitive functions can be performed with differential efficiency both over individuals and over certain other variables, such as developmental periods. Such is the case with a large number of physiological functions, notably circulation and respiration, whose fundamental automaticity is beyond question. Following up this analogy, we may hypothesize that frequency judgments are, while essentially automatic, nonetheless subject to some secondary influences, including developmental factors. The way in which respiration, more clearly than heart rate, can be consciously modified, is an analogous example. While this interpretation also does seem to be consistent with the fact that prior knowledge of a frequency test and even specific instructions to count frequencies during study typically have little effect on the accuracy of frequency judgments (Hasher & Zacks, 1984), the contradictory reports and the opposite perspectives are also presented on this issue (Greene, 1984; Fisk, 1986; Zacks, Hasher, & Hock, 1986). However, as the second implication, it is also true that results found in this study do not completely qualify the alternative of the automaticity, namely the availability hypothesis which is a direct application of the availability heuristic of Tversky and Kahneman (1973). The implication is that any factor that affects the memorability of instances of an event will have an effect on frequency estimations of that event. Availability is judged both by how many examples of a certain event can be recalled and by how quickly and easily the examples come to mind. Finally, as the third implication, it is still possible that there is a continuum linking completely automatic and nonautomatic processes and that frequency coding is not at the extremely automatic end of this continuum. Many processes might be involved both in coding frequency information at the time of the word presentation and in retrieving the information at the time of test to estimate frequencies. There may be contributions of direct and indirect mechanisms of coding that accumulate counts of frequency (Underwood, 1969) and that rely automatic and some not. Also, though each contribute to the mechanisms of coding to some extent, some may dominate under certain conditions. A finer grained analysis of the frequency judgment which permits to partition all the processes involved in generating frequency judgments would surely be needed.

As Tables 1 and 6 indicate, the superiority of the 10th graders in frequency estimation was clear and consistent, suggesting that it cannot be attributed to the test delay. The age-related variation pattern of the frequency judgment accuracy was considerably stable, particularly when the correlational analyses of the estimated and actual frequencies were made, as in Table 6. Thus, the unexpectedly large decrement in absolute frequency judgments following retention testing reported by Erickson and Gaffney (1985) may thus be tentatively interpreted as due to the greater amount of their intervening testing.

The analyses of the item-frequency variable provide some interesting and potentially valuable suggestions as to the source of the superiority in frequency judgment shown by the 10th graders. As indicated in

Table 3, that superiority was primarily evidenced on items that occurred once. A similar interaction occurred in the free-recall data, as shown in Table 4, with the superiority of the 10th graders most evident on the F1 and F3 items. For the F5 items, the like superiority pattern was weak as revealed in Table 3 or reversed as shown in Table 4. But these results need to be regarded with some caution, because no control for item specificity over event frequencies was made. Nonethless, it is difficult to imagine how any interaction of specific item and age could account for these results in lieu of the item-frequency variable.

We may further consider the implications of the age-frequency results for the interpretation of the frequency-judgment process. Apparently, the 10th graders were better able to take account of the occurrence of, and to recall the names of the items that have been only presented once. If this differential performance was based on some kind of superior cognitive alertness, it is possible that such alertness reaches an optimal level at about 16 years of age and that it, rather than the ability to detect, retain, retrieve, and report frequencies per se, can account for much of our results. While this issue needs to be further researched, it should be noted that this interpretation is more consistent with the hypothesis of automaticity and uniformity over aging that has been advanced by Hasher and Zacks (1979, 1984).

The interaction between item frequency and test delay (Table 5) also merits consideration. It was apparently due to the dispersion of scores by test delay at the F0 and F1, rather than the F3 and the F5 frequency levels. Again there is the suggestion that the higher item frequences are generally resistant to variables such as test delay, and that the differences in overall performance are primarily accounted for by differences in the ability to estimate the occurrence or the nonoccurrence of the F0 and the F1 items. Because the more frequent items were equally well estimated after a test delay up to one month (Table 5) by the four age groups (Table 3), our data may actually be interpreted as offering support for the automaticity hypothesis, on the premise that is it the large frequencies that are more crucial in the frequency-judgment process.

The differential effects of the variables in these data on the F0 an F1 items raises the question of the relationship between recognition memory and frequency judgment. Although Proctor (1977) found some differences between recognition scores and frequency judgments of one or more frequencies items grouped together scored for recognition, that result has not been replicated. The evidence indicating that recognition and frequency-judgment tests tap the same memory system and utilize essentially the same retrieval technique has been summarized by Harris, Begg, and Mitterer (1980); the results in their own experiments were consistent with those earlier reported by Malmi (1977) in showing a lowered criterion for accepting items as "old", which includes both hits and false alarms, for the frequency judgments compared to the recognition judgments, but no difference in the net discriminability of old items. More recently, Hintzman and Stern (1984) have reported data on forgetting rates of recognition and event-frequency measures that are consistent with the assumption of a common mechanism.

The commonality of free-recall and frequency-judgment functions suggests that there is some kind of underlying commonality of cognitive processing. Although specification of such processing is hardly feasible at this time, nonetheless the differential efficiency in both of the retrieval operations does seems to be dependent upon variables like age and test delay. The data in this study strongly suggest that the age differences observed in frequency judgment were due to differential recognition memory. The hypothesis is supported by analyses of the retrospective reports made by subjects immediately following a frequency-judgment test. These protocols suggesting the operation of some different kind of factors by the older subjects such as high school and college students compared to sixth graders (Marx, 1986) is entirely consistent with the age-differentiated results of the present study. However the theoretical interpretations develop, it seems clear that future research on the frequency-judgment problem should provide separate analyses of the very low event frequencies rather than simply lumping all of the frequencies together in the analysis of the data.

One other especially interesting result deserves comment. This is the finding that female subjects were superior to male subjects in frequency judgment. It was true primarily and consistently after the test delay of 7 days and one month. With respect to free recall, there was also a surprising consistency in female superiority to male subjects over the test delays, suggesting again their relative resistance to forgetting over the delay intervals. The developmental generality of these results over the age span tested is indicated by the absence of any interaction of age with gender.

REFERENCES

Attig, M.S. and Hasher, L. (1980) The processing of frequency information by adults. *Journal of Gerontology*, 35-69.

Attneave, F. (1953) Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology*, 46, 81-86.

Barsalou, L.W. and Ross, B.H. (1986) The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 116-134.

Erickson, J.R. and Gaffney, C.R. (1985) Effects of instructions, orienting task, and memory tests on memory for words and word frequency. *Bulletin of the Psychonomic Society*, 23, 377-380.

Fisk, A.D.(1986) Frequency encoding is not inevitable and is not automatic: A reply to Hasher and Zacks. *American Psychologist*, 41, 215-216.

Fisk, A.D. and Schneider, W. (1984) Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 181-197.

Ghatala, E.S. and Levin, J.R. (1973) Developmental differences in frequency judgments of words and pictures. *Journal of Experimental Child Psychology*, 16, 495-507.

Goldstein, D., Hasher, L., and Stein, D.K. (1983) The processing of occurrence rate and item information by children of different ages and abilities. *American Journal of Psychology*, 96, 229-241.

Greene, R. L. (1984) Incidental learning of event frequency. *Memory and Cognition*, 12, 90-95.

Harris, G., Begg, I., and Mitterer, J. (1980) On the relation between frequency estimates and recognition memory. *Memory and Cognition*, 12, 90-95.

Hasher, L. and Chromiak, W. (1977) The processing of frequency information: An Automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, 16, 173-184.

Hasher, L. and Zacks, R. T. (1979) Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.

Hasher, L. and Zacks, R.T. (1984) Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist*, 39, 1372-1388.

Hintzman, D.L. (1969) Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology*, 80, 139-145.

Hintzman, D.L. and Stern, L.D. (1984) A comparison of forgetting rates in frequency discrimination and recognition. *Bulletin of the Psychonomic Society*, 22, 409-412.

Howell, W.C. (1973) Storage of event frequencies: A comparison of two paradigms in memory.

Journal of Experimental Psychology, 98, 260-263.

Johnson, M.K., Raye, C.R., Hasher, L., and Chromiak, W. (1979) Are there developmental differences in reality monitoring? *Journal of Experimental Child Psychology*, 27, 120-128.

Kausler, D. H., Lichty, W., and Freund, J. S. (1985) Adult age differences in recognition memory and frequency judgments for planned versus performed activities. *Developmental Psychology*, 21, 647-654.

Kausler, D.H., Lichty, W. and Hakami, M.K. (1984) Frequency judgments for distractor items in a short-term memory task: Instructional variation and adult age differences. *Journal of Verbal Learning and Verbal Behavior*, 23, 660-668.

Kausler, D. H. and Puckett, J. M. (1980) Frequency judgments and correlated cognitive abilities in young and elderly adults. *Lournal of Gerontology*, 35, 376-382.

Lund, A. M., Hall, J.W., Wilson, K. P., and Humphreys, M. S. (1983) Frequency judgment accuracy as a function of age and school achievement(learning disabled versus non-learning-disabled) patterns. *Journal of Experimental Child Psychology*, 35, 236-247.

Maki, R. H. and Ostby, R.S. (1987) Effects of level of processing and rehearsal on frequency judgments. *Journal of Experimental Psychology: Learning, Memory and Cognion*, 13, 1, 151-163.

Malmi, R. A. (1977) Context effects in recognition memory: The frequency attribute. *Memory and Cognition*, 5, 123-130.

Marx, M.H. (1985) Retrospective reports on frequency judgments. *Bulletin of the Psychonomic Society*, 23, 309-310.

Marx, M.H. (1986) More retrospective reports on event-frequency judgments: Shift from multiple traces to strength factor with age. *Bulletin of the Psychonomic Society*, (In press)..

Naveh-Benjamin, M. and Jonides, J. (1985) The effects of rehearsal on frequency coding. *Bulletin of the Psychonomic Society*, 23, 387-390.

Naveh-Benjamin, M. and Jonides, J. (1986) On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention. *Journal of Experimental Psychology: Lenarning, Memory, and Cognition*, 12, 3, 378-386.

Proctor, R. W. (1977) The relationship of frequency judgments to recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 679-039.

Tversky, A. and Kahneman, D. (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

Underwood, B.J. (1969) Attributes of memory. *Psychological Review*, 76, 559-573.

Underwood, B.J. and Zimmerman, J. (1973) Serical retention as a function of hierarchical structure. *Journal of Experimental Psychology*, 99, 236-242.

Underwood, B.J., Zimmerman, J. and Freund, J. (1971) Retention of frequency information with observations on recognition and recall. *Journal of Experimental Psychology*, 87, 149-162.

Williams, K.W. and Durso, F. (1986) Judging category frequency: Automaticty or availability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 3, 387-396.

Zacks, R. T., Hasher, L., and Hock, H.S. (1986) Inevitability and automaticity: A response to Fisk. *American Psychologist*, 41, 216-218.

원고 초본 접수 : 1987. 4. 9
최종 수정본 접수 : 1987. 7. 30

韓國心理學會誌
Korean Journal of Psychology
1987. Vol. 6, No. 1. 21－33

事象의 頻度判斷과 自由再生의 比較 : 年齡, 性 및 檢査遲延의 効果

김 영 채 · Melvin H. Marx

계명대학교 심리학과 Georgia State University

被驗者의 年令, 性 및 檢事遲延을 函數로 하여 事象의 頻度判斷과 自由再生 反應이 어떻게 共變化 하는지를 比較해 보았다. 頻度破斷의 自動性 假說을 대부분의 認知機能은 年令과 더불어 發達해 감에도 불구하고 事象의 頻度判斷은 이러한 發達傾向을 보이지 아니하며 그것은 努作的 過程이 아니라 융통성이 없고 接近 不可能한 自動的過程이라고 주장한다. 그리고 이러한 頻度略號化의 自動性 假說은 頻度段階의 예민성은 年令에 따라 별로 變化하지 아니한다는 것을 포함한 다섯가지의 準據에 따라 檢證되고 있다. 일련의 動物이름을 하나씩 슬라이드로 제시한다음 576名의 被驗者에게 먼저 이들을 自由再生케 한다음, 이어서 각기의 動物 이름이 몇번 提示되었느냐는 頻度(0, 1, 3 또는 5회)를 推定해 보게 하였다. 年令變因(國交6年, 高校1年, 大學生, 40對成人)과 檢事遲延(0, 1, 7 또는 30일) 變因에 따른 自由再生 및 頻度判斷의 反應結果는 대단히 유사한 變化樣態를 보여 주었다. 그리고 전체적으로 보아 高校1年生 集團이 가장 우수 하였으며 修行水準은 檢事遲延에 따라 감소하였다. 이러한 結果들은 正確反應數를 準據로 하였을 때나 實際 頻度와 推定頻度間의 相關 係數를 準據로 利用 했을때도 모두 근본적으로 보아 동일하였다. 얻은 結果는 지금껏 報告된 바 없거니와 이들의 理論的意義를 頻度判斷은 生涯範疇에 걸쳐 근본적으로 一定하다는 自動性 假說에 따라 論議해 보았다.