Technical Papers

The Influence of Glutamic Acid on Test Performance[1]

D. G. Ellson, Paul R. Fuller, and Robert Urmston

Department of Psychology, Indiana University,[2] Bloomington

Zimmerman, Burgemeister, and Putnam (*3*) have recently reported an experiment from which they conclude that "glutamic acid accelerates mental functioning in human subjects," and "the most striking results appear in the seriously retarded group, where statistically significant differences are obtained between test and retest intelligence quotients." Because of a number of criticisms that may be raised against the design of this experiment and some of the statistical treatments, the study has been repeated by the authors with certain necessary changes.

The basic objection to the original experiment is the lack of adequate control for the effects of knowledge on the part of the test administrators that glutamic acid was being administered. As controls Zimmerman used 37 of the 60 subjects included in the experimental group, determining changes in their I.Q.'s during a period 6 months to 8 years prior to the experiment. Inasmuch as the experimental and control treatments occurred at different times, it seems almost certain that those who administered the tests knew that glutamic acid was being given between the pre- and post-tests. At least, no attempt to control this important variable is reported. Because the scoring of parts of the Stanford-Binet is far from objective, it is quite possible for enthusiastic testers unintentionally to modify scores in favor of the experimental results.

A second criticism of the Zimmerman experiment concerns the selection of the retarded group, which was done in such a way as to produce a regression error in favor of the obtained difference. The mean score of any subgroup that is selected on the basis of test scores, and that has a mean score different from that of the larger group, can be expected to regress toward the mean of the larger group when retested. The expected change in the subgroup mean is proportional to $D(1\text{-}r)$, where r is the correlation between pre- and post-test scores, and D is the difference between the two means. In this case, r is the reliability coefficient of the Stanford-Binet. The obtained reliability is not given, but assuming an r of .90, the estimated regression is 1.3 I.Q. points, reducing the gains due to other variables present in the experiment from 6.3 to 5.0.

A third criticism concerns the computation of D/σ_D. Sufficient data are presented in Zimmerman's Tables I and II to indicate that the correlation term was omitted in computing σ_D. The correlation involved is again the reliability of the Stanford-Binet. Conservatively estimating this r as .90, we find that D/σ_D is raised from 1.65 to 5.04 in the total group, and from 3.02 to 11.05 in the retarded group, indicating that highly significant gains were actually obtained in both groups. Unfortunately, statistically significant differences have no meaning unless experimental controls are adequate.

The present experiment was designed for more adequate control by the use of two groups run simultaneously under identical conditions except for the experimental variable. Until the experiment was completed, only one person, who did not otherwise participate in the experiment, knew which group received glutamic acid, and which the placebo.

Subjects for the experiment were two groups of 30 children, inmates of Muscatatuck State School, who ranged in age from 9 to 17 years, and in Binet I.Q. from 16 to 70. The two groups were matched by pairs for age and I.Q., as shown in Table 1.

TABLE 1

RESULTS OF MATCHING EXPERIMENTAL AND CONTROL GROUPS

Group	Binet I.Q.		Age	
	M	σ	M	σ
Experimental	49.05	13.6	12.5	2.07
Control	49.00	14.9	12.9	2.26

Four tests were administered in addition to the Stanford-Binet: The Cornell-Coxe Performance Ability Scale, a tapping rate test, a coordination test (star tracing), and a test of memory span for digits. The test-retest reliability of these measures is presented in Table 2.

TABLE 2

TEST-RETEST RELIABILITY OF SCORES

Test	9/47 to 2/48	2/48 to 6/48	9/47 to 6/48
Stanford-Binet (I.Q.)	.94	.93	.95
Cornell-Coxe	.93	.94	.95
Tapping rate	.91	.91	.86
Coordination			
Time	.67	.67	.37
Errors	.80	.74	.72
Digit span	.92	.89	.79

Tests were administered to all the subjects by psychometrists from the Department of Psychology at Indiana University in September, 1947, February, 1948, and

[1] The writers wish to acknowledge the very helpful cooperation of the administration and personnel of the Muscatatuck State School and of Benson Scodel, who provided medical supervision.

[2] This research was supported by a grant from the Graduate School of Indiana University, which provided a one-year fellowship for Robert Urmston.

TABLE 3

Mean Differences in Test Scores on Successive Test Administrations. Negative Values Indicate Poorer Performance on the Second Test of Each Pair

Test	Initial mean score	9/47 to 2/48			2/48 to 6/48			9/47 to 6/48		
		D	t	p	D	t	p	D	t	p
A. *Experimental group*										
Stanford-Binet (I.Q.)	49.05	5.00	5.59	< .01	0.18	.28	> .05	5.18	4.39	< .01
Cornell-Coxe	109.00	17.93	3.76	< .01	1.07	.18	> .05	19.00	4.61	< .01
Tapping rate*	62.50	7.25	3.90	< .01	0.00	...	> .05	7.25	4.03	< .01
Coordination										
Time	13.21	0.07	.06	> .05	1.07	1.22	> .05	1.14	1.03	> .05
Errors	24.21	2.78	1.81	> .05	− 1.28	1.04	> .05	1.50	.85	> .05
Digit span	4.65	0.15	.08	> .05	0.07	0.05	> .05	.22	.12	> .05
B. *Control group*										
		D	t	p	D	t	p	D	t	p
Stanford-Binet (I.Q.)	49.00	4.05	3.00	< .01	− 0.73	.58	> .05	3.32	2.86	< .01
Cornell-Coxe	99.33	22.00	5.74	< .01	3.54	.53	> .05	25.54	4.11	< .01
Tapping rate*	59.19	4.25	2.58	< .05	0.94	.42	> .05	5.19	2.19	< .05
Coordination										
Time	12.36	− 3.14	2.76	< .05	− 0.71	.60	> .05	− 3.85	2.12	< .05
Errors	25.00	− 6.00	3.30	< .01	1.93	1.66	> .05	7.93	4.02	< .05
Digit span	4.57	0.03	.02	> .05	0.31	0.12	> .05	0.34	0.12	> .05

* Tapping rate as here given is number of taps in 15 sec, average for 10 trials.

June, 1948. Forms L and M of the Stanford-Binet were alternated in the three testing periods, the remaining tests being the same in each period. Test administrators were the same for the first two and different for the third.

Following the first administration of the test battery, the 60 subjects were divided into two matched groups identified (to school personnel) as A and B. Group B was given sodium-neutralized 1(+) glutamic acid mixed in 8 oz whole milk or tomato juice. The initial daily dose was 4 g, increasing over a period of 42 days to the maximum of 30 g, which was maintained for the remainder of the experiment. Group A, the control group, was given a placebo, also mixed in milk or juice, consisting of 14 g corn sugar, 4 g sodium bicarbonate, and 1 g table salt, which produced a taste change in the mixture somewhat similar to the glutamic acid. The dosage was given to both groups after breakfast and at 2:30 P.M. The glutamic acid and placebo were provided in dated containers labeled only A and B. The powders were mixed with the liquid and administered to the appropriate groups by school personnel. Throughout the entire experiment every possible precaution was taken to prevent both school personnel and psychometrists from knowing which treatment was given to either group.

Table 3 presents the mean change in test scores and the statistical significance of the changes for both groups. The Stanford-Binet I.Q., Cornell-Coxe scores, and tapping rate show significant gains for both groups, with the major part of the gain occurring during the first 5 months of experimentation. The coordination test showed a significant increase in time and decrease in errors in the control group, but not in the experimental group. No significant changes were found for the digit span test.

TABLE 4

Comparison of Mean Changes in Test Scores for Experimental and Control Groups. Negative Differences Indicate That Less Improvement Was Shown by Experimental Group Than by Control Group

Test	9/47 to 2/48			2/48 to 6/48			9/47 to 6/48		
	D	t	p	D	t	p	D	t	p
Stanford-Binet (I.Q.)	.95	.63	> .05	.91	.70	> .05	1.86	1.16	> .05
Cornell-Coxe	− 4.07	.57	> .05	− 2.47	.347	> .05	− 6.54	.85	> .05
Tapping rate	3.00	1.46	> .05	− .94	.41	> .05	2.06	.71	> .05
Coordination									
Time	3.21	2.02	> .05	1.78	1.33	> .05	5.00	2.44	< .05
Errors	− 3.22	1.42	> .05	− 3.21	1.74	> .05	− 6.43	3.72	< .01
Digit span	.12	.048	> .05	− .24	.888	> .05	− .12	.0323	> .05

Table 4 presents the comparison of changes in experimental and control group mean scores. The experimental group showed slightly greater gain in I.Q. and tapping rate and smaller gain in Cornell-Coxe performance than the control group, but differences were not significant. The only significant differences between the gains made by the two groups were found in coordination test scores, where the slower and more accurate performance of the star-tracing task by the control group is significantly different from the performance of the experimental group.

For comparison with individual data presented by Zimmerman and collaborators, Table 5 presents the gains in I.Q. for the 3 individuals in each group showing maximum gain.

TABLE 5

GAINS IN I.Q. FOR THREE INDIVIDUALS IN EACH GROUP SHOWING MAXIMUM GAIN FROM 9/47 TO 6/48

Experimental group				Control group			
Subject	I.Q. 9/47	I.Q. 6/48	Gain	Subject	I.Q. 9/47	I.Q. 6/48	Gain
Em	74	95	21	Ma	57	76	19
Bu	51	62	11	Mo	51	68	17
Cu	44	55	11	El	63	75	12

The results presented above provide strong support for the hypothesis that controls in Zimmerman's experiment were inadequate. Our control group showed significant gains in I.Q. that did not differ significantly from those in the experimental group. The present experiment gives no evidence concerning the cause of these gains, but it may be suggested that a bias in test scoring, which was apparently not controlled in the earlier experiment, may have been operative. In the Indiana study another variable may be mentioned. The teachers at the school were greatly interested in the experiment. Consequently, the stereotyped institutional life of the subjects was very probably livened up by extra attention. The subjects in both groups also made two trips a day to the school kitchen. This additional stimulation may have produced some effect on test scores. Zimmerman and his collaborators do not report their procedure in sufficient detail to indicate whether extra stimulation was present in their study.

The mean gain in I.Q. reported for the retarded group in the Zimmerman study was greater than in the present one. However, when gains in the retarded group are corrected for regession as estimated, they are approximately the same as were found in the Indiana experimental group. It seems likely that, if bias affects test scores, the effect would be greater in the Zimmerman experiment provided we can assume that the psychometrists knew that the subjects were receiving glutamic acid. In the Indiana experiment the psychometrists knew that only half the subjects were receiving this treatment.

Although extreme gains in individual scores are not acceptable evidence when variable measures are used, it is of interest, though not surprising, that gains comparable to those reported in the appendix to Zimmerman's study were also found in the Indiana study (Table 5). Comparable extreme gains appear in both the experimental and the control groups.

The significant difference in time taken for star-tracing in the coordination test is in apparent agreement with animal studies (*1, 2*), which have shown greater activity following administration of glutamic acid. The significantly poorer error performance of the experimental group is probably related to the difference in time.

If tapping rate may be considered a measure of maximum rate of response, it is of interest that this was not significantly different in the two groups.

References

1. HAMILTON, H. C., and MAHER, E. B. *J. comp. physiol. Psychol.*, 1947, **40**, 463.
2. MARX, M. H. *J. comp. physiol. Psychol.*, 1949, **42**, 313.
3. ZIMMERMAN, F. T., BURGEMEISTER, B. B., and PUTNAM, T. J. *Psychosom. Med.*, 1947, **9**, 175.

The Measurement of Radioactive Hydrogen in Solid Samples—Comparison with Gas Counting[1]

Maxwell Leigh Eidinoff and Joseph E. Knoll

Department of Physics and Biophysics, The Sloan-Kettering Institute for Cancer Research,* and *Department of Chemistry, Queens College, New York City

Radioactive hydrogen has been measured as hydrogen gas, or a volatile compound containing hydrogen, inside Geiger-Müller counters and ionization chambers (*1, 2, 4, 6*). The experiments reported below demonstrate that solid compounds, suitably plated out on a sample pan and using a windowless counter, can be measured with satisfactory precision. The principal disadvantage of the solid-counting procedure arises from the extremely weak energy of the β particles emitted by radioactive hydrogen, leading to an infinite thickness of less than one mg/cm². The most important advantage is the elimination of the time-consuming steps in the gas-counting measurement: the quantitative combustion of the sample and complete conversion[2] of the resulting water into hydrogen gas. Additional time is consumed in the preliminary "seasoning" of the combustion-conversion train to avoid isotopic contamination. Further, the compound is not destroyed during the analysis. Consequently, it is expected that solid-counting techniques will be a valuable complementary tool in radioactive hydrogen tracer studies, including paper chromatographic and radioautograph techniques.

The compound selected for study of solid-counting was methyl-3-α-acetoxycholinate containing radioactive hydrogen in the 11 and 12 positions. This compound, obtained in the course of a steroid tracer project, was made by the hydrogenation of the $\Delta^{11,\,12}$ cholenate in an acetic acid medium containing radioactive hydrogen, with the use of a platinum oxide catalyst. It was purified by repeated addition and removal of inactive solvent and was recrystallized three times from petroleum ether (60° C). The melting point was 134.5° C.

The solid was directly plated on aluminum pans having an area of approximately 10 cm². A solution of the

[1] This project is jointly supported by the Office of Naval Research Contract #N6-ori-99, T.O.1., and the Atomic Energy Commission.

[2] Incomplete conversion results in isotopic fractionation and leads to incorrect analyses.