*THE EFFECT ON THE BEHAVIOR OF CHILDREN,
AS REFLECTED IN THE IQ SCORES,
WHEN REINFORCED AFTER EACH CORRECT RESPONSE*

CALVIN V. EDLUND

SAN JUAN UNIFIED SCHOOL DISTRICT

This experiment studied the effect on intelligence test scores of a probable reinforcer given for correct responses. Eleven pairs of 5- to 7-yr-old children were matched on the basis of a strong liking of candy, no physical problems associated with eating it, parent permission to receive and eat the candy, age, sex, and a revised Stanford-Binet Scale Form L IQ score. The control group was given the revised Stanford-Binet Scale Form M, as prescribed in the test manual. The experimental group was also given Form M according to the manual, except M&M candy was given for each plus or correct response. There was an appreciable, statistically significant difference between the resulting IQ test scores of the two groups.

Several studies have sought to determine the effect various rewards have on children's intelligence test scores. Hurlock (1924, 1925) studied the effects of praise and blame in three early experiments. In the first two, she used the National Group Intelligence Tests. In the first experiment, she divided third-, fifth-, and eighth-grade children into groups by age, sex, race, and intelligence, and on a test-retest design, over a one-week span, studied the effects of praise and blame. After the first testing, one group was praised for its test scores, and a second group was reproved for its test scores. The second experiment was a followup of the first and in both she found no difference between the effects of praise and blame, yet both were more effective than practice alone.

Benton (1936) explored the effect on children's IQ scores of praise for their original intelligence test performance and the promise of prizes if they did better on the second. He matched his seventh- and eighth-grade subjects for age, sex, IQ score, and grade. He found that a control group lacking these incentives differed insignificantly from the experimental group.

Klugman (1944) also studied the effects of the promise of two presumed reinforcers on IQ scores. He made use of alternate forms of the Stanford Binet, 1937 revision. His subjects were 72 children from grades 2 to 7. Praise and money were the reinforcers given if the children did better on the second testing. Although he found no significant difference overall between the two reinforcers, scores of the Negro children were substantially raised when tested with money as compared to praise, and scores of the white children were raised more by praise than were the scores of the Negroes.

The effects of reinforcement on the IQ scores of 480 second and third graders divided into three groups (middle-class white, lower-class white, and lower-class Negro) were studied by Tiber and Kennedy (1963). A 1960 revision of the Stanford Binet was used. Experimental incentives were given at the end of each subtest. Four reinforcement groups were distinguished: praise, reproof, candy, and no reinforcement. No significant differences among groups were found.

Considering the studies reviewed above, there would seem to be several important factors worthy of further study. In all of these investigations, the comparative effectiveness of several reinforcers was studied. In none of these studies

¹Reprints may be obtained from the author, San Juan Unified School District, 3738 Walnut Avenue, Carmichael, Calif., 95608.

did the experimenters use a known effective reinforcer in terms of improving the rate of occurrence of the desired behavior (correct responses on an IQ test) for all subjects. Further, there is no indication that any of the studies considered the deprivation state of the subject relative to the chosen reinforcers. Yet, the Klugman study showed the importance of using the appropriate stimulus for the child under study.

The precision of reinforcement needs additional consideration. Of the studies reviewed, some subjects received reinforcers for responding or for total IQ improvement. It would be far more precise to reinforce only correct responses, since an IQ score is the result of many such responses and is raised as the number of correct responses is accumulated.

In some of the studies, reinforcement was delivered on a delayed schedule: *i.e.*, they were promised reinforcement if they did better.

The purpose of the present investigation was to study the effect of a presumptive reinforcer, chosen on the basis of its high probability of being effective, given contingent on correct responses to an individual intelligence test, as evidenced in the total IQ score.

METHOD

Subjects

The 79 children from low-middle class and lower-class homes used in this study were from 5 to 7 yr of age. All were attending public school, Headstart, and Kindergarten classes.

Procedure

The subjects were given the revised Stanford Binet, Form L Test. In addition, it was determined whether or not each child liked candy, by individually asking the child and his parents. The parents were also consulted to determine whether or not their child had problems digesting candy. Twenty two, 11 pairs, were found to match on the basis of intelligence test scores, age, sex, a liking of candy, no problem digesting candy, and parent permission to take the candy. The matched pairs included 10 pairs of boys and

one pair of girls. At random, one subject from each pair was assigned to an experimental group and the other then assigned to a control group.

The experimental group was given Form M of the revised Stanford Binet Test seven weeks after the alternate form, Form L, was given. In administering Form M, this departure from the test manual was made: the subject was told: "I am going to give you an M&M candy for each right answer you give to the questions I ask and each thing you do right that I ask you to do." Subsequently, the subject was immediately given an M&M candy for each correct performance. No candy was given when a minus response was made by the subject. The control group was also given Form M in seven weeks. The test was given according to the instructions in the manual (Terman 1937). All examinations were given just before lunch, so their food-deprivation state was at a high level.

Experienced school psychologists conducted the testing. To hold constant the effect of the examiner, Form M was administered by the same person.

The revised Stanford Binet Scales, Forms L and M, were used to obtain the intelligence test scores of all the children in this study. These were chosen for their similarity of measurement throughout and because "they are mutually equivalent in difficulty . . ." (Terman 1937).

A *t* test for matched pairs was used to evaluate the resultant mean scores.

RESULTS

The IQ scores of each subject are indicated in Table I. The subjects in the experimental group, with one exception, improved; the control group showed little improvement, five showing poorer scores on the second testing. The median gain for the experimental group was 12 IQ points, as compared to a median gain of one IQ point for the control group.

The means of experimental and control groups on the previous Stanford Binet, Form L Test were identical, 82, and the standard devia-

Table I

Revised Stanford-Binet Scales, L & M, I.Q. scores of experimental and control groups.

Pairs	Experimental Group			Control Group		
	L	M	D*	L	M	D*
a	75	85	10	71	76	5
b	71	83	12	75	82	7
c	78	97	19	79	76	-3
d	66	79	13	61	68	7
e	100	119	19	111	96	-15
f	94	106	12	91	85	-6
g	87	107	20	81	86	5
h	61	72	11	74	73	-1
i	83	90	7	79	90	11
j	85	95	10	84	85	1
k	107	107	0	98	97	-1

*D indicates the difference between first and second scores.

tion scores were 14.18 and 13.67 respectively. These scores indicate that the original pairing on the basis of IQ scores was satisfactory.

A *t* test of the difference between the means of the matched experimental and control groups on the Stanford Binet, Form M, IQ scores, yielded a *t* of 3.92, with 10 degrees of freedom. This was significant at the 0.01 level of confidence. A summary of the results of the *t* test is presented in Table II.

Table II

Summary of t-Test for Means of Stanford-Binet Form M Intelligence Quotients Between Experimental and Control Groups.

Group	Mean I.Q.	df	<i>t</i>	<i>t</i> 0.05	<i>t</i> 0.01
Experimental	94	10	3.92	2.23	3.17
Control	83				

DISCUSSION

The results indicate that a carefully chosen consequence, candy, given contingent on each occurrence of correct responses to an IQ test, can result in a significantly higher IQ score. The subjects receiving reinforcers for each correct response, with one exception, had much higher rates of correct responses than those who did not receive reinforcers.

This study would suggest that environmental factors, more specifically the reinforcement con-

ditions holding between testee and tester, need to be seriously considered in interpreting the IQ concept. Such a model would suggest the IQ is a numerical way of coding and summarizing several behaviors an organism has learned, and will perform if motivated.

The immediate practical results of the present study would be for those using IQ scores for administrative purposes. It would seem important that precise reinforcement procedures need to be used in the testing procedure if one is to produce an accurate summary of the individual's learning progress or his IQ, which may be used as a basis for administrative decisions.

It may prove informative and interesting to pursue this line of study further, using reinforcers other than those used thus far. It may be that other reinforcers may have an even more marked effect on the subject's test behavior and, thus, his IQ. This could add to our understanding of the magnitude of the environmental contribution to intelligence and provide further practical direction to those who are engaged in testing humans.

REFERENCES

- Benton, A. L. Influence of incentives upon intelligence test scores of school children. *Journal of Genetics and Psychol.* 1936, **49**, 494-497.
- Edlund, C. V. *How human behavior is learned.* Seattle: Special Child Publications. 1971.
- Hurlock, Elizabeth B. The value of praise and reward as incentives for children. 1924. *Archives of Psychology*, No. 71.
- Hurlock, Elizabeth B. An evaluation of certain incentives used in school work. *Journal of Educational Psychology*, 1925, Vol ?? 145-159. (b)
- Klugman, S. F. The effect of money incentives versus praise upon the reliability and obtained scores of the Revised Stanford-Binet Test. *Journal of Genetics and Psychology* 1944, **30**, 255-269.
- Terman, L. H. and Merrill, M. A. *Measuring Intelligence.* 1937. Houghton, Mifflin, New York.
- Tiber, N. and Kennedy, W. A. The effects of incentives on the intelligence test performance of different social groups. *Journal of Consulting Psychology*, 1964, **28** (2), 187.

Received 15 April 1969.

(Revised 18 February 1972.)