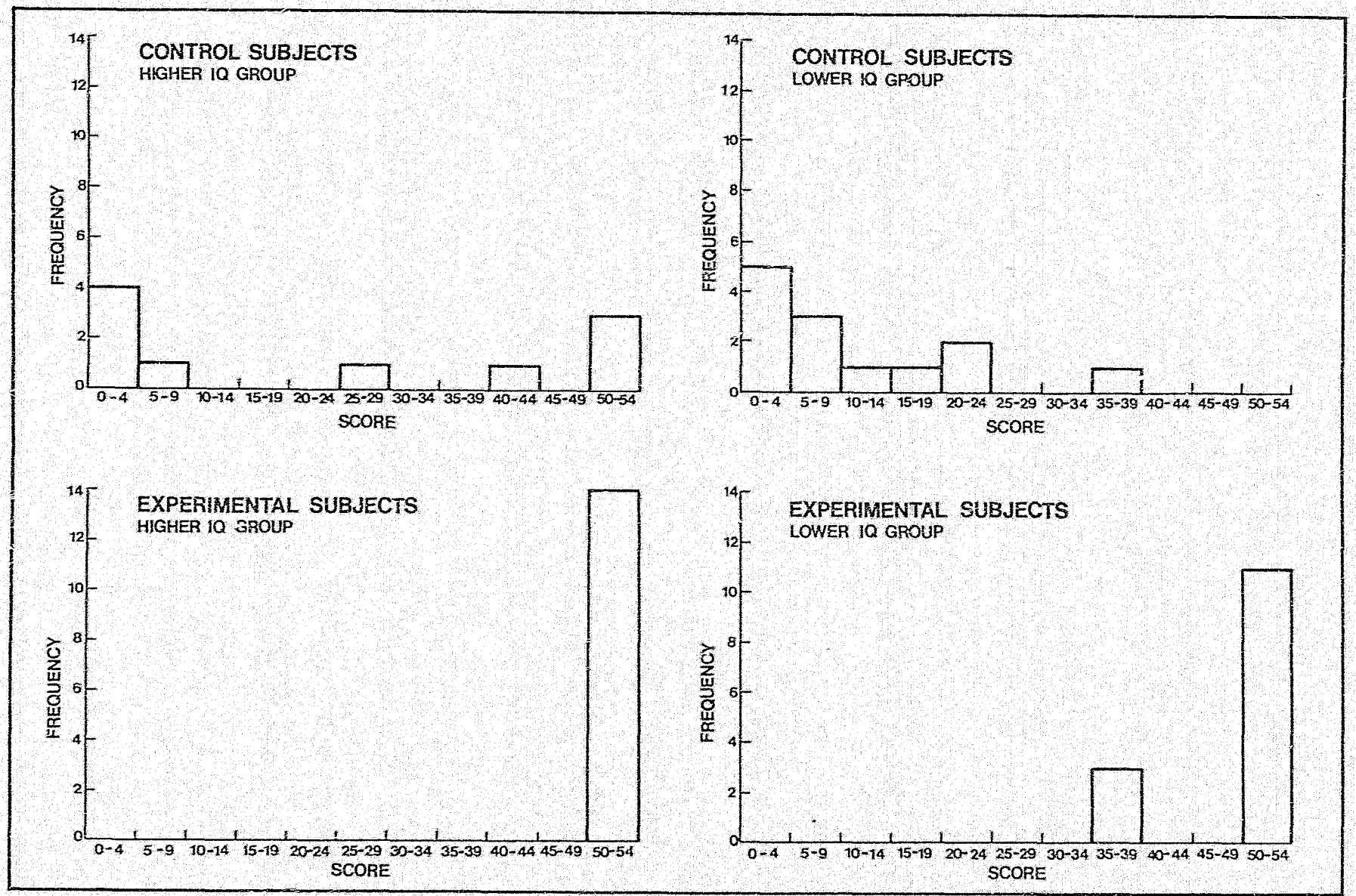

Figure 10.1 Scores Achieved on Producing a Factorial Array for Rates of Chemical Reaction by Standard Nine Boys

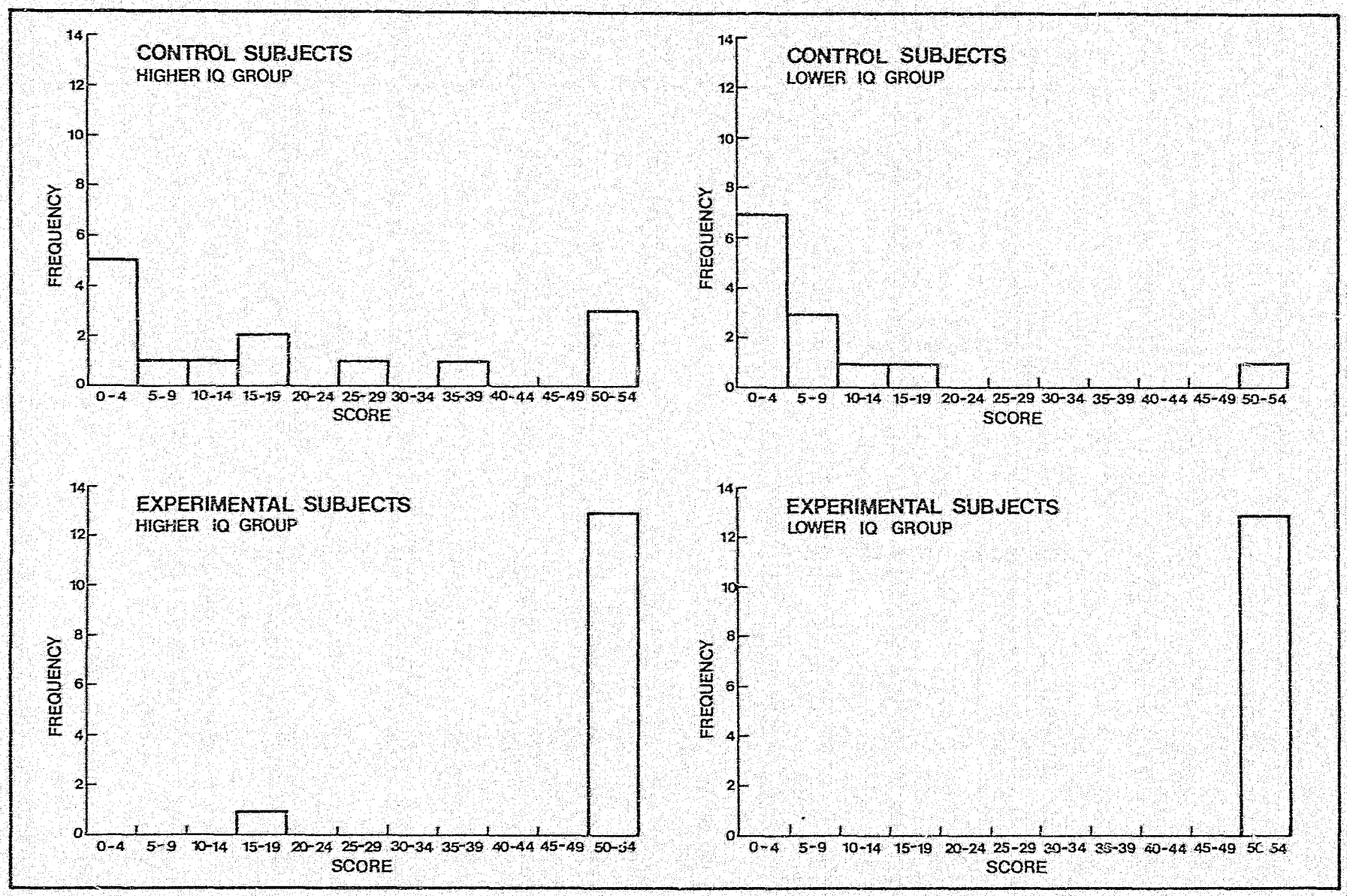

Figure 10.2 Scores Achieved on Producing a Factorial Array for Rates of Chemical Reaction by Standard Eight Boys

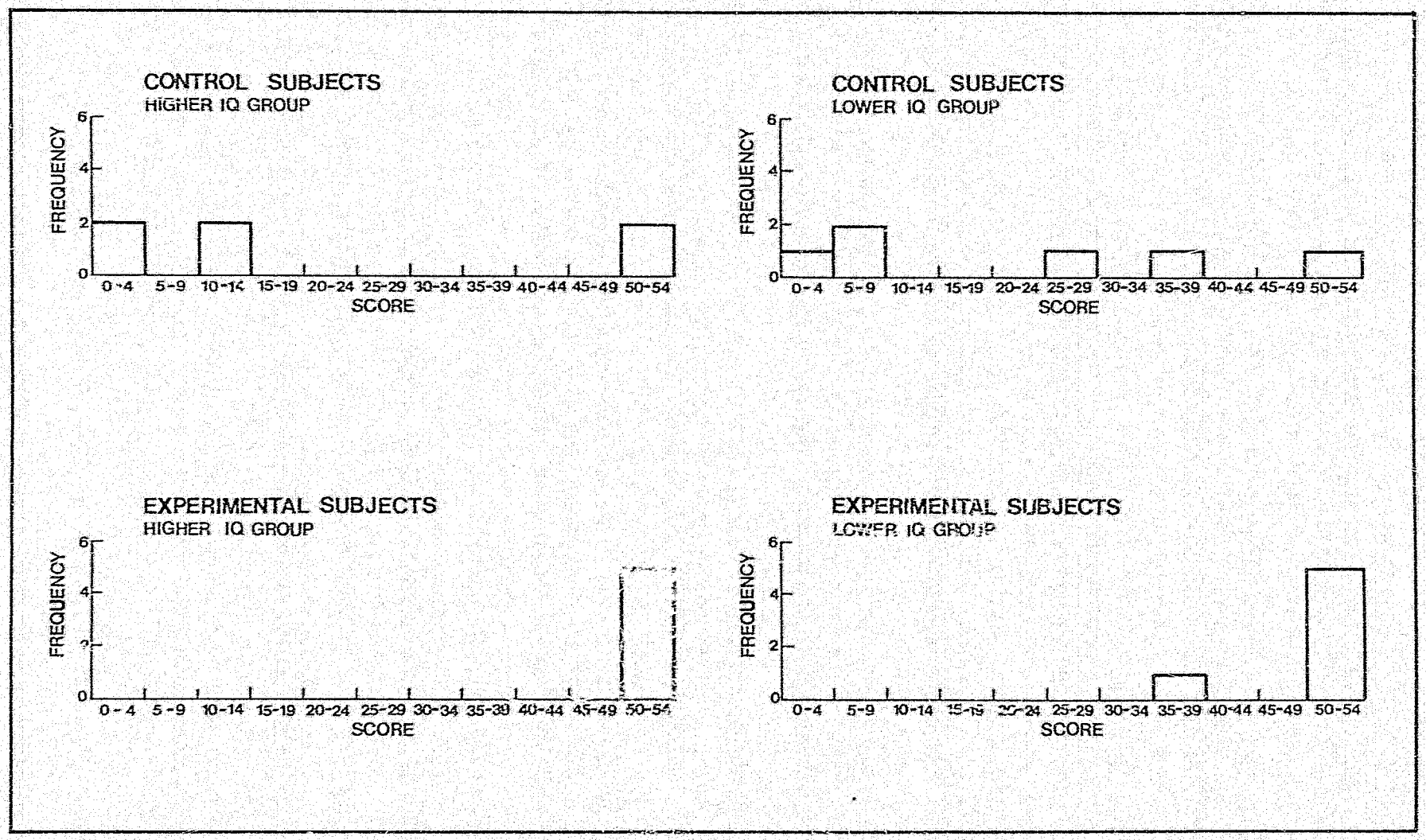

Figure 10.3 Scores Achieved on Producing a Factorial Array for Rates of Chemical Reaction by Standard Nine Girls

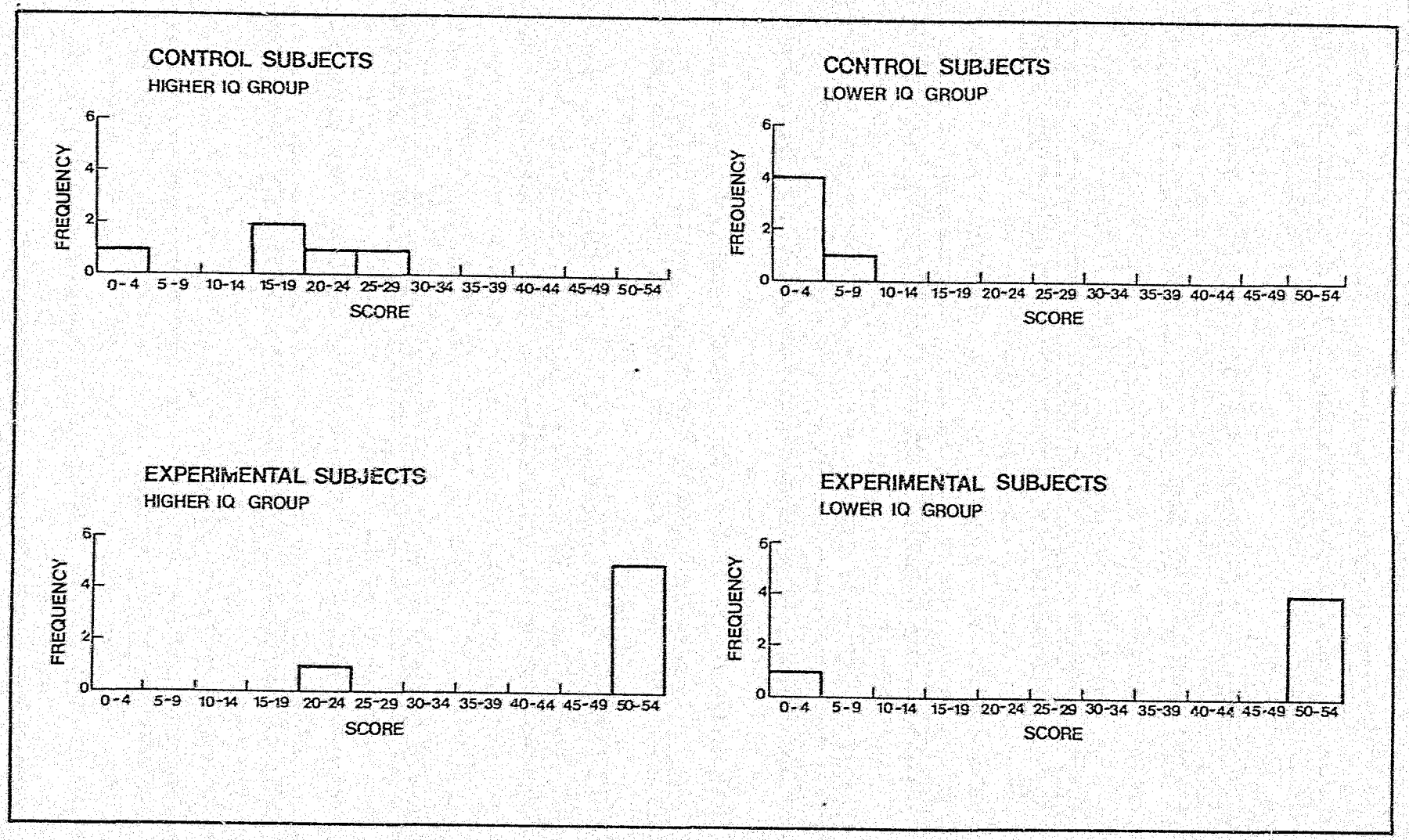

Figure 10.4 Scores Achieved on Producing a Factorial Array for Rates of Chemical Reaction by Standard Eight Girls

All the members of the experimental group supplied tree diagrams and the overall correctness of these indicated that the training was successfully transferred to the novel learning material. Among the combined experimental group, five pupils clearly understood the work but omitted all combinations which involved no catalyst. They had thus failed to consider that the absence of something constituted a level and could lead to additional combinations. Two other members of this group made careless errors of omission in their haste to complete their tree diagrams. Another pupil produced a tree diagram which did not contain a single combination which was correctly or even meaningfully represented. The subject with the highest IQ, a Standard Nine boy, responde gebraic symbols. Instead of supplying a tree expressed in words as shown i raining, he supplied a key with which to interpret his time-efficient symbolically expressed tree. It is noteworthy that all control subjects who achieved high scores used a pattern or semi-pattern with which to generate combinations. Nine control subjects scored full scores of fifty-four points. All of these had employed block patterns of some kind which, although not identical to the pattern contained in the tree diagram, were systematically similar. All nine were higher IQ subjects and the possibility of IQ-dependence in this test item is therefore examined as well as the research hypothesis. Such IQ-dependence is not expected, however, as the logically similar tasks of Piaget's first chemical experiment and the train problem did not show any marked IQ effects.

Table 10.1 shows that the scores achieved by the experimental group were significantly higher than those achieved by the control subjects in every block. Statistical analysis therefore confirms the qualitative observations above, namely, that transfer of the training by the experimental group resulted in superior achievement by this group. In two of the lower IQ blocks, the observed values of the Mann-Whitney test statistic were extremely low, suggesting that in some cases the lower IQ subjects might have benefited markedly from training. Further, as mentioned above, the nine control subjects who obtained full scores belonged to higher IQ groups. The question of IQ-dependence is therefore examined in Table 10.2. Table 10.2 indicates that the effect of IQ is significant only in the case of the control block of Standard 8 girls. In five of the other seven blocks, however, the observed value of U was drawn from the higher IQ group. In the control block of Standard 9 girls, U came from the lower IQ group by a marginal difference between the two groups. In the experimental block of Standard 8 boys, U was the same whether drawn from the higher or lower IQ groups. There thus seems to be a tendency, although not significant, for higher IQ groups to perform better.

Table 10.1 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.1 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

Since the proportion of ties in each case was large, the test statistic was corrected for ties using the normal approximation (Siegel, 1956), except where the observed value of U was very much smaller than the critical value of U.

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_0
	Experimental Group	Control Group	Experimental Group	Control Group	Observed	Critical, $\alpha=0,05$			
Boys									
High IQ Std.9	14	10	22,5	117,5	22,5	41	-3,31	0,0005	reject
Low IQ Std.9	14	13	1,5	180,5	1,5	56	--	--	reject
High IQ Std.8	14	14	25,5	170,5	25,5	61	-3,69	0,0001	reject
Low IQ Std.8	13	13	1	168	1	51	--	--	reject
Girls									
High IQ Std.9	5	5	5	25	5	--	-2,12	0,0170	reject
Low IQ Std.9	5	6	4,5	31,5	4,5	--	-2,31	0,0104	reject
High IQ Std.8	6	5	1	29	1	25,5	--	0,004	reject
Low IQ Std.8	5	5	4,5	20,5	4,5	--	-1,74	0,0414	reject

Decision : Reject H_0 in favour of H_1 : the experimental subjects performed significantly better than the control subjects.

Table 10.2 Rank Ordering of Scores on the Generation of a Factorial Array : Examination of IQ Effects

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in all experimental group blocks since the proportion of ties was large and in the single control group block which was a borderline case, using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_0
	High IQ	Low IQ	High IQ	Low IQ	Observed	Critical, $\alpha=0,05$			
Boys									
Std. 9 Expt.	14	14	79	121	75	61	1,58	0,0571	accept
Std. 9 Control	10	13	47,5	82,5	47,5	37	--	--	accept
Std. 8 Expt.	14	13	91	91	91	56	0,00	0,5000	accept
Std. 8 Control	14	13	60,5	121,5	60,5	56	-1,48	0,0694	accept
Girls									
Std. 9 Expt.	5	6	12,5	17,5	12,5	--	-0,91	0,1814	accept
Std. 9 Control	6	6	19	17	17	--	--	0,469	accept
Std. 8 Expt.	6	5	14	16	14	--	-0,27	0,3936	accept
Std. 8 Control	5	5	1,5	23,5	1,5	--	--	<0,016	reject

whether trained or untrained. It will be remembered that the same tendency was observed in responses to both Piaget's first chemical experiment and in the train task.

10.6.2 Recall of Knowledge

The responses to the test item which required producing a combination of factors which could result in an explosive reaction, are analysed in Table 10.3. In this and in other items of the post-test, except where otherwise specified, a one-tailed test has been used, even though the training did not involve such items. If the null hypothesis were to be rejected, the direction of the difference is predicted in favour of the experimental group, for reasons discussed earlier in this chapter.

Table 10.3 shows that the null hypothesis is accepted for every block. While the differences between experimental and control groups are not significant, the observed value of the Mann-Whitney test statistic was drawn from the experimental group for six of the eight blocks, including all the blocks of boys. This indicates a weak tendency for the trained subjects to achieve higher scores.

Table 10.4 involves recall of the terminology relevant to rates of reaction. The table shows that the experimental group obtained better scores than the control group in seven of the eight blocks but the superior performances reached statistical significance in only two of the blocks, namely, the Standard 9 boys of both higher and lower IQ. In each of these two blocks, the position of the observed value of z is not far from the rejection region, so there is unlikely to be a very marked dissimilarity in performance between the trained and untrained subjects. Performance differences, overall, were therefore only marginally in favour of the experimental group. Because of this, the independent factors of age, IQ and sex were examined to test whether these variables are more important predictors of success in learning than any acceleration due to the training.

Table 10.3 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.2 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in the borderline case using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_0
	Experimental Group	Control Group	Experimental Group	Control Group	Observed	Critical, $\alpha=0,05$			
Boys									
High IQ Std.9	14	10	45	95	45	41	-1,58	0,0571	accept
Low IQ Std.9	14	13	75	107	75	56	–	–	accept
High IQ Std.8	14	14	88,5	107,5	88,5	61	–	–	accept
Low IQ Std.8	13	13	62,5	106,5	62,5	51	–	–	accept
Girls									
High IQ Std.9	5	6	12,5	17,5	12,5	–	–	> 0,331	accept
Low IQ Std.9	6	6	22	14	14	–	–	0,294	accept
High IQ Std.8	6	5	17,5	12,5	12,5	–	–	> 0,331	accept
Low IQ Std.8	5	5	9,5	15,5	9,5	–	–	> 0,274	accept

Table 10.4 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.3 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in borderline cases or where the proportion of ties was large using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_0
	Experimental Group	Control Group	Experimental Group	Control Group	Observed	Critical, $\alpha=0,05$			
Boys									
High IQ Std.9	14	10	41	99	41	41	-1,77	0,0384	reject
Low IQ Std.9	14	13	55	127	55	56	-1,77	0,0384	reject
High IQ Std.8	14	14	87,5	108,5	87,5	61	-0,49	0,3121	accept
Low IQ Std.8	13	13	65	104	65	51	-1,00	0,1587	accept
Girls									
High IQ Std.9	5	6	8,5	21,5	8,5	–	–	> 0,123	accept
Low IQ Std.9	6	6	21	15	15	–	–	0,350	accept
High IQ Std.8	6	5	13	17	13	–	–	0,396	accept
Low IQ Std.8	5	5	9	16	9	–	–	0,275	accept

The effects of age are considered in Table 10.5 which shows that the older subjects in three of the blocks performed significantly better. In all eight blocks, however, the observed value of U was drawn from the Standard 9 group, which was indicative of superior performances by older subjects in general. This result may in fact underestimate age effects since an earlier conclusion in the thesis was that division of the sample by standard did not give sufficient separation of ages to enable resolution in assessment of performances. It is noteworthy that this topic is normally introduced at Standard 10 level, a practice recommended by the present finding of increasing mastery with age.

The influence of IQ is examined in Table 10.6 which indicates that the higher IQ subjects performed better in all blocks, except for the control block of Standard 9 girls. This generally superior achievement reached statistical significance in two of the blocks, namely, the experimental block of Standard 9 boys and the control block of Standard 8 boys. There is thus a tendency for higher IQ subjects to be more successful in learning the topic. The moderate IQ-dependence found, lends some degree of support to the tenet that ability, as measured by a standardised IQ instrument, is among the most useful indicators of science learning in general (Boulanger, 1981; Butts, 1980; Lynch *et al*, 1979). Results were not so strongly IQ-dependent in a study by Lynch and Dick (1980), which investigated a specific aspect of science learning, namely, the recognition of definitions of science concepts such as *solid, liquid, gas* and *mole*. These workers showed that secondary school pupils of high IQ (125+) were subject to very erratic performance with grade compared with the relatively smooth development in aggregate performance of other pupils (high IQ sample, $n_1 = 63$ and other sample, $n_2 = 1635$). They observed a slight underperformance by the high IQ group at Grade 7 but, by Grade 10, the high IQ group achieved significantly better on nine out of the sixteen concept definitions.

Table 10.7 shows that there was no significant difference between the scores achieved by the boys and by the girls, although U was drawn from the girls in all eight cases, indicating that gender-bias tended to favour female subjects. Personal teaching experience suggests that, although boys may often understand the work better, the conscientiousness of many female pupils frequently enables girls to surpass the boys in school science examinations. Research findings mostly show that sex is not an unfailing predictor of success in science instruction. Maccoby and Jacklin (1974), in their comprehensive review of sex differences, conclude that, during adolescence, boys attain greater mastery of scientific knowledge. Butts (1981), on

Table 10.5 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.4 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in borderline cases using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_o
	Std. 9	Std. 8	Std. 9	Std. 8	Observed	Critical, $\alpha=0,05$			
Boys									
High IQ Expt.	14	14	56,5	139,5	56,5	61	−1,96	0,0250	reject
High IQ Control	10	14	68	72	68	41	–	–	accept
Low IQ Expt.	14	13	57,5	124,5	57,5	56	−1,64	0,0505	accept
Low IQ Control	13	13	52,5	116,5	52,5	51	−1,65	0,0493	reject
Girls									
High IQ Expt.	5	6	11,5	18,5	11,5	–	–	> 0,268	accept
High IQ Control	6	5	15	15	15	–	–	0,535	accept
Low IQ Expt.	6	5	13,5	16,5	13,5	–	–	> 0,396	accept
Low IQ Control	6	5	6	24	6	–	–	0,0368*	reject

*The exact distribution of U in the case of ties was determined as described by Lehmann (1975, p. 19).

Table 10.6 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.5 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in borderline cases using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_o
	High IQ	Low IQ	High IQ	Low IQ	Observed	Critical, $\alpha=0,05$			
Boys									
Std. 9 Expt.	14	14	57,5	138,5	57,5	61	−1,91	0,0281	reject
Std. 9 Control	10	13	51	79	51	37	–	–	accept
Std. 8 Expt.	14	13	62,5	119,5	62,5	56	−1,39	0,0823	accept
Std. 8 Control	14	13	43,5	138,5	43,5	56	–	–	reject
Girls									
Std. 9 Expt.	5	6	6,5	23,5	6,5	–	–	0,0649*	accept
Std. 9 Control	6	6	18,5	17,5	17,5	–	–	>0,469	accept
Std. 8 Expt.	6	5	14	16	14	–	–	0,465	accept
Std. 8 Control	5	5	8	17	8	–	–	0,210	accept

*The exact distribution of U in the case of ties was determined as described by Lehmann (1975, p.19).

the other hand, states in his review article that young adolescent boys and girls can be expected to have similar achievement outcomes with some exceptions where girls achieve better than boys. Lynch and Paterson (1980) have dealt with the effects of sex in the specific issue of recognition of science concept definitions. Their samples were the same ones studied by Lynch and Dick (1980) above and spanned the four years of high school. For groups of pupils similarly matched in terms of IQ, they report large numbers of significant differences on the sixteen concept terms. All these differences except one favoured the boys. A similar study by Erickson and Erickson (1984) lends support to the findings of Lynch and Paterson.

10.6.3 Application of Knowledge

The test items requiring application of knowledge sometimes elicited mere rephrasing of the questions as a response. This phenomenon was most marked in Question 13 (Section C, in Appendix G). In answer to Question 13, forty pupils from the combined sample of experimental and control subjects, reworded the question. Most of these were Standard 8 pupils. It was not clear to the examiner whether such responses indicated lack of understanding or miscomprehension.

The overall scores achieved on Section C are analysed in Table 10.8. The null hypothesis is accepted in every block. In four of the blocks, the experimental group tended to achieve better; in three, the control subjects tended to achieve better, while in the remaining block, there was no detectable difference between the groups. It is therefore concluded that the training did not enhance achievement where searching questions requiring insight into the topic were posed.

The possibility that IQ is the dominant predictor of examination performance is investigated in Table 10.9. In six of the blocks, higher IQ subjects achieved better and significantly surpassed their lower IQ counterparts in three of these blocks. In the remaining two blocks, the lower IQ subjects tended to achieve better, namely, the control block of Standard 8 girls and the experimental block of Standard 8 boys. For each of these blocks, however, the Mann-Whitney statistics were very similar for both higher and lower IQ subjects, indicating little difference. The conclusion from Table 10.9 is that IQ has a marked tendency to influence achievement.

Table 10.7 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.6 under Test

Mann-Whitney U Test, two-tailed, $\alpha = 0.05$

The test statistic was corrected for ties in borderline cases using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_0
	Boys	Girls	Boys	Girls	Observed	Critical, $\alpha = 0,05$			
Standard 9									
High IQ Expt.	14	5	38	32	32	13	–	–	accept
High IQ Control	10	6	37,5	22,5	22,5	11	–	–	accept
Low IQ Expt.	14	6	43	41	41	17	–	–	accept
Low IQ Control	13	6	60,5	17,5	17,5	16	-1,91	0,0562	accept
Standard 8									
High IQ Expt.	14	6	51	33	33	17	–	–	accept
High IQ Control	14	5	41	29	29	13	–	–	accept
Low IQ Expt.	13	5	44,5	20,5	20,5	12	–	–	accept
Low IQ Control	13	5	48	17	17	12	-1,54	0,1236	accept

Table 10.8 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.7 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for the ties in borderline cases using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_0
	Experimental Group	Control Group	Experimental Group	Control Group	Observed	Critical, $\alpha = 0,05$			
Boys									
High IQ Std.9	14	10	65	75	65	41	–	–	accept
Low IQ Std.9	14	13	68	114	68	56	-1,13	0,1292	accept
High IQ Std.8	14	14	117	79	79	61	–	–	accept
Low IQ Std.8	13	13	59	110	59	51	-1,32	0,0934	accept
Girls									
High IQ Std.9	5	6	17,5	12,5	12,5	–	–	> 0,331	accept
Low IQ Std.9	6	6	24,5	11,5	11,5	–	–	> 0,155	accept
High IQ Std.8	6	5	11	19	11	–	–	0,268	accept
Low IQ Std.8	5	5	12,5	12,5	12,5	–	–	> 0,500	accept

Table 10.9 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.8 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in borderline cases using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U		z	p	Decision re H_o
	High IQ	Low IQ	High IQ	Low IQ	Observed	Critical, $\alpha = 0,05$			
Boys									
Std. 9 Expt.	14	14	89,5	106,5	89,5	61	–	–	accept
Std. 9 Control	10	13	53	77	53	37	–	–	accept
Std. 8 Expt.	14	13	95	87	87	56	–	–	accept
Std. 8 Control	14	13	49,5	132,5	49,5	56	–2,03	0,0212	reject
Girls									
Std. 9 Expt.	5	6	3,5	26,5	3,5	–	–2,13	0,0166	reject
Std. 9 Control	6	6	6,5	29,5	6,5	–	–1,86	0,0314	reject
Std. 8 Expt.	6	5	14,5	15,5	14,5	–	–	< 0,465	accept
Std. 8 Control	5	5	14	11	11	–	–	0,421	accept

10.6.4 Compensation Skills

Responses to the test item involving compensation skills, revealed that only six experimental subjects and five control subjects realised that, within the context of the given question, more than one combination existed which would give a suitable rate of reaction. Many responses also displayed evidence of the misconception mentioned earlier, namely, the erroneous relationship between particle size and surface area. This misconception had also been detected in the answers to Section B, Question 9(ii), Appendix G (combination resulting in an explosive reaction) and in Section C, Question 11 (applications question involving surface area). Responses to all these questions were evaluated in terms of this misconception. Table 10.10 shows that, although most pupils held the correct conception, a considerable proportion of the pupils would require remedial teaching.

Table 10.10 Conceptions of the Relationship between Particle Size and Surface Area

Conception	Number of Pupils	
	Experimental Group	Control Group
Correct relationship	52	39
Incorrect relationship	18	25
(Assessment not clear)	7	8

Statistical analysis of compensation reasoning abilities as manifested on the test item is given in Table 10.11. A two-sided test was used since the training was not expected to influence skills other than combinatorial reasoning. Ausubel (1964) has emphasised that

> *If stages of development have any true meaning, ... although some acceleration is possible, it is necessarily limited in extent.* (Ausubel, 1964, p. 264).

Table 10.11 shows that the experimental subjects obtained higher scores in five blocks, reaching statistical significance in the case of the lower IQ Standard 9 boys. In the other blocks, the control subjects achieved higher scores with significance in one block, namely, the lower IQ Standard 9 girls. It is concluded, therefore, that the training did not extend to include enhancement of compensation skills as measured by the test item.

Table 10.11 Rank Ordering of Scores on the Posttest on the Novel Instructional Material : Null Hypothesis 10.9 under Test

Mann-Whitney U Test, two-tailed, $\alpha = 0,05$

The test statistic was corrected for ties in borderline cases using the normal approximation (Siegel, 1956).

Block	Sample Size		Mann-Whitney Statistic		Value of U				
	Experimental Group	Control Group	Experimental Group	Control Group	Observed	Critical, $\alpha = 0,05$	z	p	Decision re H_o
Boys									
High IQ Std. 9	14	10	64	76	64	36	–	–	accept
Low IQ Std. 9	14	13	46	136	46	50	-2,21	0,0272	reject
High IQ Std. 8	14	14	109	87	87	55	–	–	accept
Low IQ Std. 8	13	13	65	104	65	45	–	–	accept
Girls									
High IQ Std. 9	5	6	16	14	14	–	–	0,930	accept
Low IQ Std. 9	6	6	35	1	1	–	–	0,004	reject
High IQ Std. 8	6	5	11,5	18,5	11,5	–	–	> 0,536	accept
Low IQ Std. 8	5	5	7,5	17,5	7,5	–	–	> 0.310	accept

10.6.5 Efficacy of the Advance Organizer

The stated aim of the thesis is to contribute towards increased achievement of pupils in physical science and both Piagetian strategies and Ausubelian strategies were therefore used in the instruction. In view of the research objective, the individual effect of each of these measures on mastery of the learning material was not considered. The components of the training in its dual role of an attempt to accelerate Piagetian developmental level and of an advance organizer for novel academic matter are intertwined and both involved in the achievement outcomes of the instruction.

The instruction was presented in an Ausubelian sequence but this does not have the corollary that the training acted as an advance organizer for the whole topic. The training functioned as advance organizer for only the particular subsumer, the generation of a factorial array for rates of reactions. However, since the training

caused not only significant gains on the factorial array which represented the concept underlying the topic but also marginal gains on the recall of factual knowledge, the advance organizer/training to a small extent benefited mastery of the subject-matter as a whole. Since the examination was by no means exhaustive, it is not clear whether the experimental group in fact understood the factual material better or whether they had more interest, motivation or confidence due to the training.

The subsumer specified in the previous paragraph also functioned as a test of transfer of the training. Such a transfer test had the advantage that it did not constitute a mere extension of the training, which was a criticism levelled by Barratt (1975) at the transfer measures used in some training studies. The test of transfer involved a novel situation which was also more complicated than the conceptual framework presented in the training. The possibility of rote learning exists, although this possibility appears to be slight as approximately three weeks had elapsed since reinforcement of the training. To correctly apply a memorised solution, a subject would in any case have to recognise the structural equivalence of the training framework and the criterion problem in the instruction and thence adapt the training to suit the more complex demands of the problem at hand.

10.6.6 Retention Test

Table 10.12 analyses the performances by experimental subjects on the train task as observed in the immediate and delayed tests. Scores on the delayed test were assigned as for the immediate test. Although the direction of any differences due to forgetting would favour the immediate test, the influence of the following factors was not known :

(i) Reinforcement of the training which was received after the immediate test.

(ii) Examination conditions compared with the interview situation of the immediate test.

(iii) Any practice effect associated with the factorial array dealing with rates of reaction earlier in the examination.

A two-sided test was therefore employed.

Table 10.12 Rank Ordering of Scores on the Train Evaluation Task as Retention Test : Null Hypothesis 10.10 under Test (Experimental Group)

Mann-Whitney U Test, two-tailed, $\alpha = 0,05$

The retention test involved trained subjects only and the proportion of ties on the upper limit of the measurement scale was large. In every case, therefore, the test statistic was corrected for ties using the normal approximation. (Siegel, 1956).

Block	Sample Size	Sum of Ranks		Observed Value of U	z	p	Decision re H_0
		Immediate Test	Delayed Test				
Boys							
High IQ Std. 9	14	182	224	77	–1,80	0,0718	accept
Low IQ Std. 9	14	231	175	70	–2,12	0,0340	reject
High IQ Std. 8	14	210	196	91	–0,60	0,5486	accept
Low IQ Std. 8	13	170,5	180,5	79,5	–0,37	0,7114	accept
Girls							
High IQ Std. 9	5	25	30	10	–1,00	0,3174	accept
Low IQ Std. 9	6	39	39	18	0,00	1,0000	accept
High IQ Std. 8	6	39	39	18	0,00	1,0000	accept
Low IQ Std. 8	6	36	42	15	–0,64	0,5222	accept

Table 10.12 shows that the null hypothesis is accepted in seven of the eight blocks. In four of these, performance had improved with time, approaching significance in the block of higher IQ Standard 9 boys. There was no detectable difference between the immediate and delayed tests in two other blocks. In the block of lower IQ Standard 9 boys, there was significant deterioration in performance with time. In the evaluation of responses to the delayed test, a considerable number of experimental subjects were penalised due to careless mistakes such as omitting some of the terminal branches of the tree and confusing which switches were two-way operational or three-way operational. Had the memory task involved manipulations as in the immediate test, such errors would necessarily have been rectified. The results shown in Table 10.12 give rise to the overall conclusion that the effects of the training were lasting.

The final point of interest was the extent to which the trained subjects had maintained their lead on the untrained subjects. Table 10.13 indicates that the experimental group was markedly superior to the control group on the delayed test. The observed value of the Mann-Whitney test statistic was extreme in five blocks and gave rise to very small probabilities in all eight cases. In four of the blocks, the values of U were not quite as extreme as on the immediate test (Table 6.1) but these cannot be compared directly for two reasons. Firstly, the number of subjects was slightly different owing to absentees and school leavers. Secondly, the scoring to obtain the data in Table 10.13 was the same as for the rates of reaction (Section 10.6.1), that is, dependent on only the number of combinations with a penalty for redundant combinations. This approach to scoring was used for simplicity and convenience as the null hypothesis at hand, unlike the previous null hypothesis, was not central to the thesis. With the more complex Piagetian-type scoring procedure employed in testing the previous null hypothesis, the degree of systematisation used by a subject represented a major contribution to the score. In the absence of such a contribution, which characterised the newly induced skills of the experimental group, the present scoring reduced the scores of this group. Taking a fairly extreme example, one experimental subject systematically produced thirty-six combinations with no redundant combinations on the delayed test. Scoring on the basis of the number of combinations, the points allocation was thirty-six. Using the scaling of the immediate test, however, the score would have been eighty.* Thus the fact that the values of U in testing the present null hypothesis were not as extreme as before, does not mean that the experimental group did not exhibit the same degree of superiority in performance over the control group as previously.

*The rank position of this subject within the sample has clearly changed. This point emphasises that, when comparing the research of different workers, the measurement scales need to be similar for meaningful conclusions to be drawn.

Table 10.13 Rank Ordering of Scores on the Train Evaluation Task as Delayed Test : Null Hypothesis 10.11 under Test

Mann-Whitney U Test, one-tailed, $\alpha = 0,05$

Since the proportion of ties was large, the test statistic was corrected for ties in every case, using the normal approximation (Siegel, 1956), except where the observed value of U was extreme.

Block	Sample Size		Mann-Whitney Statistic		Value of U				Decision re H_0
	Experimental Group	Control Group	Experimental Group	Control Group	Observed	Critical, $\alpha=0,05$			
Boys									
High IQ Std. 9	14	10	0	140	0	41	–	–	reject
Low IQ Std. 9	14	13	2	180	2	56	–	–	reject
High IQ Std. 8	14	14	28,5	167,5	28,5	–	−3,41	<0,0003	reject
Low IQ Std. 8	13	13	11,5	157,5	11,5	–	−3,90	0,00005	reject
Girls									
High IQ Std. 9	5	6	2,5	27,5	2,5	–	−2,50	0,006?	reject
Low IQ Std. 9	6	6	0	36	0	7	–	0,001	reject
High IQ Std. 8	6	5	0	30	0	5	–	0,002	reject
Low IQ Std. 8	6*	5	1	29	1	5	–	0,004	reject

*This group contains one more subject than the previous tables in this chapter, since one of the pupils was present for only the memory test section of the examination.

Table 10.14 shows that the majority of control subjects had now adopted approaches which were at least partially systematic in their attempts to generate a factorial array. A proportion of control subjects still approached the task in a completely random fashion : 18,1% as compared with 18,5% on the immediate test. Previously, no control subject had succeeded in using a pattern effectively. On the delayed test, however, not only did many use a pattern throughout their attempts, but three control subjects were enabled to achieve full scores. The search patterns used by members of the control group have been analysed in Chapter 6 and will not be repeated here.

Table 10.14 Degree of System Used by Control Subjects

System	Percentage of Control Subjects
Random	18,1
Pattern which disintegrated	56,9
Block pattern throughout	11,1
Other pattern throughout	11,1
(Misunderstood question)	2,8

Factors which could explain why patterns were more effective for the control subjects on the delayed test were :

(i) Communication with experimental subjects.
(ii) Enforced paper record.
(iii) Test-retest learning on the train task.
(iv) Practice effect from the factorial array involving rates of reaction, required earlier in the examination.
(v) Substitution of examination conditions for the interview situation.
(vi) Absence of laboratory equipment which required manipulation and might cause distraction from mental activities as suggested by Pallrand (1979).

Most of these points also apply to the experimental group as mentioned previously.

10.6.7 Summary of Statistical Results*

1. Experimental subjects scored significantly better than control subjects on the concept underlying rates of chemical reactions.

2. Higher IQ groups, whether trained or untrained, tended to perform better on the above concept but this effect was not significant except in the block of Standard 8 female control subjects.

*See the footnote at the end of Chapter Four.

3. Experimental subjects tended to achieve higher scores than control subjects on recall of factual knowledge but their superior performance reached significance only in the case of the Standard 9 boys on the terminology item.

4. Older subjects tended to recall terminology better but this was significant only in three of the blocks.

5. Higher IQ pupils also tended to perform better on this item, reaching significance in two of the blocks.

6. There was no significant difference between the scores achieved by the boys and by the girls, although performance favoured female subjects in all cases.

7. Trained subjects did not achieve any better than untrained subjects on the application of knowledge.

8. Higher IQ subjects, whether trained or untrained, tended to perform better on the above item. This effect reached significance for all Standard 9 girls and Standard 8 male control subjects.

9. Trained subjects did not demonstrate significantly better compensation skills than untrained subjects.

10. There was no significant difference in general between the scores achieved by the experimental group on immediate and delayed tests on the train task.

11. Experimental subjects in all blocks achieved significantly better than control subjects on the delayed test on the train task.

CHAPTER 11

CONCLUSIONS AND RECOMMENDATIONS

11.1 The Central Focus

Most of the science concepts, their extensions and derivatives that are taught in schools are imbedded in the logical structure of a mature discipline. The present literature survey has shown that acceleration of intellectual development is the preferred method to aid in the understanding of such concepts by pupils. It is believed that, if the ability to reason formally can be improved, then general achievement in school science performance will follow.

The thesis consisted of an experimental training study which aimed to induce combinatorial reasoning in adolescents. The training was intrinsically Piagetian but acted simultaneously as an operationally defined Ausubelian ad[illegible] organizer for novel learning material. Both Piagetian and Ausubelian tenets [illegible] herefore used as complementary tools to promote achievement in physical science. The exact extent of overlap of the two theories depends on the stance of the viewer with respect to, *inter alia,* discipline, psychological and sociological pressures. The fact that contributors to research in this common area often manifest contrasting interpretations, aims, methods and identify heterogeneous sets of problems, is paradoxically reassuring. An overall convergence of opinion as to the degree of interrelation of the theories of Piaget and Ausubel would indicate a facile approach to the highly complex domain of human learning.

The work in this thesis was divided broadly into four sections :

(i) The reporting of a procedure for trouble-free administration of Piaget's first chemical experi[illegible]t, in answer to such requirements as expressed in the literature;

(ii) Analysis of the structure of Piaget's first chemical experiment, in order to contribute towards the need for consistent standards in evaluating formal combinatorial thought by means of this task;

(iii) The development and execution of a training procedure aimed at enhancement of Piagetian cognitive levels;

(iv) The utilisation of the training procedure as an Ausubelian advance organizer for novel academic matter.

These four partly overlapping groups of problems constituted the scope of a systematically unified study. The main emphasis of the thesis was directed towards questions falling into the last two of the above questions.

11.2 Experimental Outcomes

11.2.1 Attainment of Objectives

The main objective of the thesis was acceleration of intellectual development to enhance achievement in physical science. Two post-training issues were therefore examined, firstly, promotion of cognitive level and, secondly, improved achievement on school learning material. The study arrived at affirmative answers to its two central questions :

1. Is it possible to develop a procedure to promote intellectual growth? In particular, is it possible to induce formal thought in the form of combinatorial reasoning which is retained and can be transferred?
2. Is it possible that the above procedure, which meets Piagetian criteria, also meets the requirements of an Ausubelian approach in the sense that it can facilitate the learning of novel academic material?

This conclusion was derived from testing the hypotheses which had been formulated at the beginning of the study in order to investigate the following issues :

(i) Prior to any intervention, were the experimental and control groups equivalent in combinatorial reasoning abilities?

(ii) After training, was the experimental group superior to the control group on appropriate tasks? Was superiority

- retained?
- transferable?

(iii) Prior to instruction, were the experimental and control groups equally unfamiliar with the selected academic material?

(iv) After teaching, was the experimental group superior to the control group

- in reproducing facts?
- in deeper conceptual insights?
- in thinking combinatorially?
- in thi[illegible]ing formally?

The decisions on the first and third of the above groups of hypotheses satisfied the statistical requirements of the research design. In the former case, the experimental and control subjects in all the blocks except one were equivalent in combinatorial reasoning abilities. The exception did not, however, in any way jeopardise the study. In the latter case, the experimental and control subjects in all the blocks were equally unfamiliar with the selected academic material.

The observed experimental outcomes of the second group of hypotheses showed that many pupils were effectively able to function and tackle selected combinatorial problems on the formal level, which they had not been able to do prior to training. This was shown, *inter alia*, by performance on the train task which was devised as an instrument which measured combinatorial skills but was not refined to the status of a Piagetian equivalent (Chapter 6). For training assessment purposes, the measurement scale reflected only degree of systematisation with penalties for redundant combinations. Trained subjects achieved significantly higher scores on the train task than untrained subjects.

Trained subjects also performed significantly better on the posttest (Chapter 7). Prior to training, the developmental levels of the experimental group as observed on Piaget's first chemical experiment were 24% concrete, 41% transitional, 24% early formal and 11% late formal. After training, these levels had improved to 1% concrete, 14% transitional, 80% early formal and 5% late formal. Training was, however, detrimental to most late formal subjects.

Specific transfer, namely, transfer to other tasks requiring combinatorial reasoning, was satisfactory since the two criterion tasks differed from the training task. A delayed test on the train evaluation task, three weeks after reinforcement of the training, sh[illegible]wed no significant difference from the immediate test (Chapter 10).

Successful training may be interpreted as induced progression along the continuum of concrete/abstract thought in that the characteristics of formal reasoning are increasingly empirically exhibited. Rote learning cannot be supposed in the presence of the criteria of generalisability and retention and thus some beneficial alteration to the cognitive structures must have occurred, even if the nature of such alterations has not been identified.

The observed experimental outcomes of the fourth group of hypotheses showed that the training, in its dual function as instrument of cognitive acceleration and Ausubelian advance organizer affected certain aspects of the material on rates of reaction to a significant extent. After teaching, the experimental group was superior to the control group

- in reproducing facts
- in thinking combinatorially

IQ, rather than training, contributed to performance on issues requiring insight into the subject-matter. There was thus no significant difference between the performances of trained and untrained subjects on items requiring application of knowledge. Higher IQ subjects tended to achieve higher scores in this area but their superiority was not generally significant. Trained subjects were not able to effect nonspecific transfer of their training, that is, they did not show any significant difference in performance from control subjects in thinking formally, as measured by a test item which required the use of compensation skills. The benefits of training observed on the learning topic as a whole were small but measurable. Any gains in achievement are worthwhile, since many pupils were thus enabled to grasp concepts with an abstract connotation. Particularly for candidates on the pass-fail border in school achievement terms, biasing their results towards the pass side by means of training may improve their motivation.

11.2.2 Effects of Independent Variables

The influences of age, IQ and sex were examined at each stage of the investigation with, *inter alia*, the major objective of establishing the impact of these variables on implementation of the training in practice. Examination of the effects of these parameters involved