This is the Preprint of H.A. Cohen and D.G. Green, *Evaluation of the Cognitive Goals of Oznaki: Enhancement of Spatial Projective Abilities*, in :"ACM Topics in Instructional Computing", A.M. Wildberger and R.G. Montanelli. eds., SIGCUE (Special Interest Group in Computers in Education) of the ACM (Association for Computer Machinery, New York, 1978, pp 69-90.

3.0 THE PROJECTION MODULE

The Projection Module developed by the authors comprises just eight lessons, whlch are detailed an the Appendix

ZONKY and WHAM involve NAKIS for which the basic commands FORWARD, BACK, RIGHT, and LEFT relate to the NAKl's current heading - not to the student's egocentric frame. To direct these NAKIs inherently requires projection.

Following Piaget's dictum, "... spatial concepts are internalized actions" in the first lessons with ZONKY pupils are encouraged to play out NAKI moves as directed by another pupil. Included in the lessons are the tasks of guiding ZONKY past obstacles, and a similar, but more demanding maze with the WHAM NAKI. The pupils are asked to punch in as many commands as possible before telling the NAKI to "Do-It".

In WHAM, students are encouraged to first use direct commands to draw some simple figure, then to work from the screen figure to write down the definition of an X (or Y, V, W) that would draw it. It is, of course, the discreteness of WHAM that makes this procedure so natural, and the lack of a record of commands is very much an advantage.

The final lessons are on a microcomputer "Life", which is an enhanced version of that described by Cohen (1978a). The "Life" NAKI obeys the commands N for North (Up), S for South (Down), E for East (Right), W for West (Left). On the command L the cell visited by the NAKI is made live, while K kills that cell. Students are asked to copy onto the screen a patterns of live cells, visible as asterisks, before initiating the Life process. (Martin Gardiner, 1970).

Of course, more than projection is taught in the lessons. They also include an introduction to the algorithmic aspects of OZNAKI, including WHAM calculator maths, some geometry, music, and a brief but purposeful introduction to some problem-solving ideas.

It is generally envisaged that the Projection globule will involve one or two students using a single OZNAKI system within an open class-room. or a complete class of say 24 students with one Wizard's Box for every two students. During the evaluation study between four and six children were taught by a single tutor (D.G.G.), using just one microcomputer Wizard Box.

4.0 EVALUATION PROGRAMME

How does one evaluate the educational robotics developed by the OZNAKI Project? Essentially there are two (intersecting) forms of evaluation. Typically educational "systems" (broadly conceived) are evaluated on a behaviorist basis: the student is conceived as some sort of black box, of unknown and unknowable inner structure. and the difference in behavior (equated to class marks and the like) is measured before and after exposure to the system. In contrast, the OZNAKI Project is concerned with the difficult task of defining cognitive goals. To specify a cognitive objective, we must state a set of changes we want to bring about in the student's

cognitive processes. Thus we are inherently concerned with the inner structuring of knowledge - where the "inner" is the contents of the behaviorist's black box.

In more familiar language, the OZNAKI Project is far more concerned with understanding than merely with **performance** of mathematical tasks. We intend to use thorough-going studies to determine whether such understanding exists. Thus we are concerned not only with the **answer** given to a particular problem, but also with the mental algorithm or **process** used to derive that answer.

It is in this general philosophic framework that studies of the OZNAKI Project were conducted in 1977. In the 1977 evaluation study we sought to examine enhancement of spatial abilities. He focussed on projection - the ability to "see" with one's mind's eye from elsewhere - this having been demonstrated by Piaget as a major process in spatial thinking.

The first trials of OZNAKI involved teaching children from Nillumbik Co-operative School. Evaluation at that stage was essentially on anecdotal data, although critical decisions about the lesson material were made. In September 1977 our first full-fledged field trials commenced in four state schools near La Trobe University. The experimental plan involved Piagetian interviews, multi-choice questionnaires, and statistical analyses of data - the subjects being students 8 - 13 years old. The Projection Module provides a general introduction to OZNAKI mathematics, with heavy emphasis on spatial skills involving projection.

4.1 EVALUATION OF THE PROJECTION NODULE

The evaluation study carried out in Spring 1977 had the following phases:

(a) Pre-testing of all students involved.
(b) Selection of Experimental and Control Groups.
(c) Teaching the Projection module to the E Group.
(d) Post-testing of both E and C Groups.

The students participating in the study were aged 8 to 13, that is at the stage of development in which projective skills are assimilated. They included both primary and secondary students.

Similar, but distinct, experiments were carried out at the high school and primary schools. Experiment 1 comprised work carried out at three primary schools: Yarraleen, Preston East, and Fawkner North. Twenty-seven pupils in seven classes completed the course. Each experimental group received two l-hour lessons per week for four weeks. Experiment 2 comprised work carried out at Montmorency High School. Forty-four pupils were involved. The 22 students in the experimental group received one l-hour lesson per week for eight weeks. Experiments 1 and 2 also had slightly different testing procedures (see below). D.G.G. conducted the lessons for all classes.

In both experiments, control groups consisted of students paired with individuals in the experimental groups. The pairs were formed by matching together students whose

performances during pre-testing were as close to identical as possible. In every case the pairs consisted of students from the same class and of similar age. We tried to reduce the effects of the children's normal learning environment (teacher, school, etc.) by selecting our subjects from as many different classes (and schools in Experiment 1) as possible. Primary school subjects were chosen at random from the entire classes available, while the secondary school subjects were all selected at random from the weakest 25% of students (as determined by the pre-test), because of our interest in the use of the Projection Module in remedial education.

Students in the control groups attended normal classes at the times when OZNAKI lessors were conducted. The effect of the course on the experimental group's ability to project was assessed by comparing their performance in post-test exercises requiring projection With the performance of their counterparts in the control group.

4.2 DESCRIPTION OF THE TESTING PROCEDURE

At both the primary and high schools, all the students involved were tested both before and after presentation of the course of lessons. Testing of the primary school children involved a Piagetian type interview, whilst testing of the high school students was conducted by written, multi-choice questionnaires.

EXPERIMENT 1

Three games were presented to the primary school children in their interviews. The same games were used in both the pre- and post-tests.

Game 1 "The mountains"

A model of three mountains was placed before the child. The interviewer produced sketches of various views of these mountains, and asked the child to place a toy man at the spot on the model from where each view could be seen. To obtain a performance parameter, the interviewer noted the time taken in each placement. Pupils scored one mark each time they placed the model man in the correct position. The Mountains is a classic test due to Piaget and his collaborators (Piaget et at, 1956, Piaget et at, 1960). In our presentation, subjects were permitted to walk around the model. Clearly the subject could solve this task by projection. However by standing behind the toy man the subject could gain nearly the same perspective.

Game 2 "Remote -Driving"

This test was designed by H.A.C. as a test whose solution projection, and everyday language and experience. The subjects were confronted with a large town plan (bare of buildings) with a rectangular grid of roads, and a "match-box" car. The children were asked to: (1) follow instructions (turn left or right) given by the interviewer, while "driving" the car about the streets of the town, keeping to the left of the centre line - this was scored on whether they made correct turns and whether they kept the car on the correct side of the road; (2) After being shown a particular route around the streets of the town plan, the subjects gave the interviewer instructions for following the route by car (that is, how many blocks to go ahead and directions of turns) one mark was scored for each correct answer. Both parts of the test demanded that the students work out directions for the car to turn for all possible orientations of the car relative to the subject.

Game 3 T he "Left Handed Route"
This test involves the use of the model town and toy car described above. In this game, the child was asked to plan a path requiring left turns only between two given points in the town and to "drive" the toy car around this path - pupils were scored simply as successful, if they found such a path, or as failed otherwise. This test could be described as a problem of planning in a spatial context.

EXPERIMENT 2

At the high school, to select the students to participate in the course, and a matched control group, we tested four complete seventh form classes using an ACER "Spatial Abilities" Test. Experimental and control group pairs were selected from the students with lowest scores on the test, by matching pupils who had both the same score and made similar sorts of errors. As in experiment 1, all pupils paired together came from the same class.

Pre-Test/post-Test: "ACER Spatial Abilities Test"
The Australian Council for Education Research (ACER.) has developed a question bank containing a collection of multi-choice answer questions involving spatial abilities mixed variously with elements of problem solving (Cornish,1977). For instance, one outline sketch shows two footprints at skew angles, and the question is whether these are two left feet, two right feet, or otherwise. We prepared two printed questionnaires both of twenty questions from these ACER. questions. One ACER. derived questionable was used with high school students both as the pretest and in selecting experimental and control group pairs. The other questionnaire was included in the post-test for high school students.

Post-Test: "The Mountains", "Remote Driving"/'The Left Handed Route"
The high school students were subjected in the post-tests to isomorphs of the problems posed to the primary children: that is, they were given exactly the same questions as the primary pupils, but in a written form instead of orally, and they marked their answers on diagrams on the test paper. The printed version of the Mountains Game was closer to the classic Planet interview in that the three model mountains were given distinctive colours.

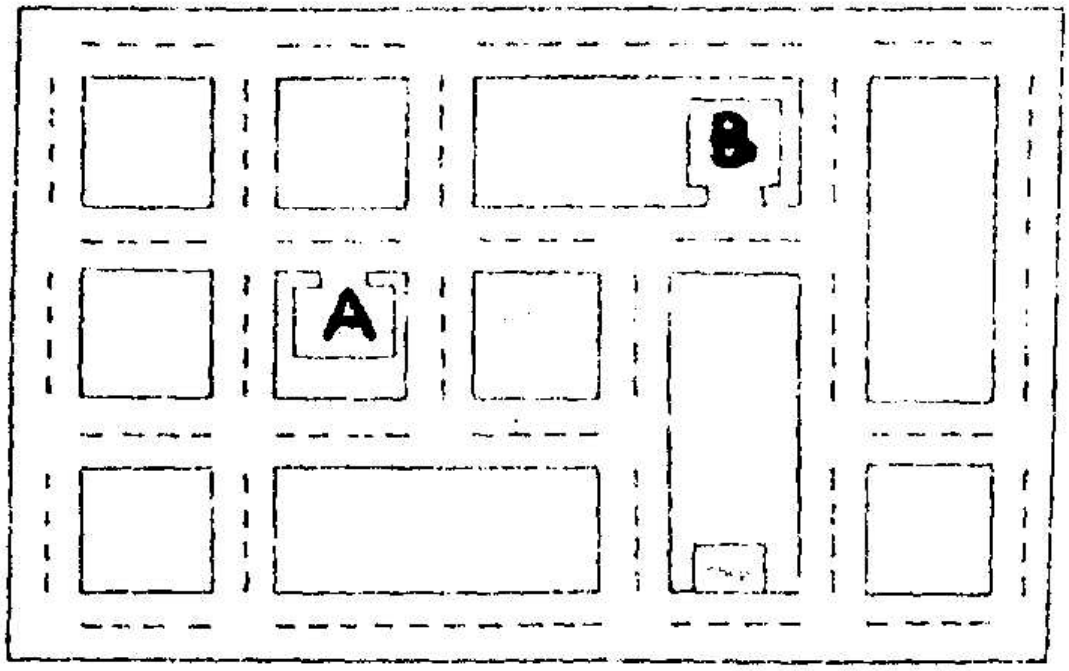

Figure 2. The town map used in the High School experiment (reduced in size). Subjects were asked to mark the path they would take in driving a car from car park A to car park B making left-hand turns only.

4.3 TEST RESULTS

Game 1 - "The Mountains"

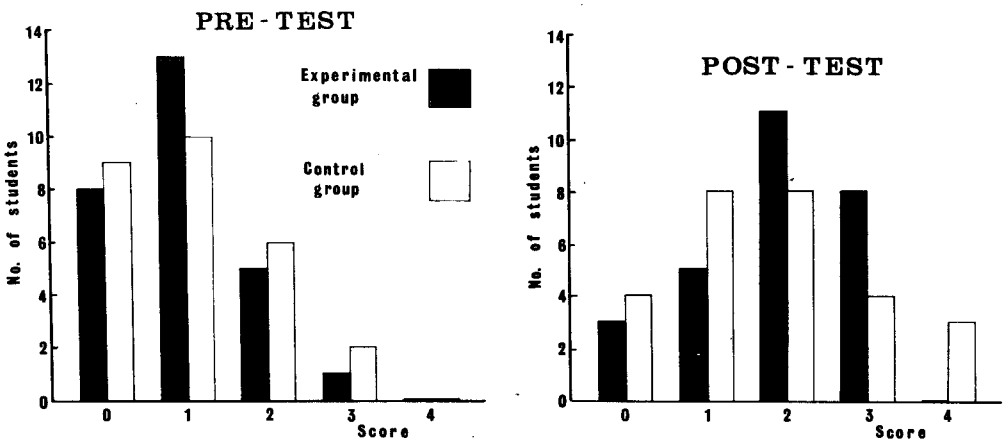

Figure 3. Score distributions in the "Mountains" Test for the primary school pupils.

Game 2 - "Remote Driving"

Table 1 presents means and standard deviations for the score distributions of the experimental and control groups in both pra- and post-testing. One-way analyses of variance were performed to test the significance of the differences in scores (Anderson, 1958; Nie et al. 1975). The probabilities resulting from an F test are tabulated: they show the probability that the differences in recorded scores did NOT occur by chance.

TABLE 1
Analysis of Pupils' Mean Scores on Remote Driving Test

	SCORES			
	Max Possible	E Group	C Group	p
Pre-test	21	12.3 (2.7)	13.1 (3.0)	.656
Post-test	21	17.9 (2.6)	14.7 (3.9)	.999
P		1.090	.916	

Notes. l. There were 27 pupils per group.
2 Numbers in brackets are standard deviations of the score distributions about the given mean value.
3. Values tabulated for : are the probabilities that the differences between the given means did NOT occur by chance.
4. A two-way analysis of variance on the results in the table showed that the significance of the interaction group and test effects was 1.00.

Game 3 - "Left-Handed Route"
Table 2 shows the numbers of pupils in each group who solved the "Left-Handed Route" problem in pre- and post-tests. The improvement shown by pupils in the experimental group was so huge that to test its significance is pointless.

The average times taken to solve the problem (by the successful students) in pre-testing were 2. minutes 36 seconds for the four experimental group students and 1 minute 41 seconds for the four control group students. In post-testing, the four experimental group students who solved the problem initially all improved their times (average l minute 30 seconds). Only one of the four control group students who solved the problem during pre-existing could again solve it during post-testing.

TABLE 2
Numbers of Pupils Solving the "Left-Handed Route" Problem

	Experimental Group	Control Group
Pre-test	4	4
Post-test	21	3

Note: There were 27 pupils in both groups.

EXPERIMENT 2 HIGH SCHOOL

Table 3 below lists means and standard deviations in the same way as for game 2 above. One-way analyses of variance were again performed on the data and the resulting significance levels for the differences in experimental and control group scores are tabulated. The two ACER space tests used were not of identical difficulty,

so the pre- and post-test scores of each group are not strictly comparable (so only one analysis of variance has been performed.) On the "Left-Handed Route" problem, 10 experimental group students and 8 control group students were successful.

TABLE 3
Analysis of Scores on High School Tests

		SCORES		
TEST	Max Possible Score	E GROUP	F GROUP	p
Pre-test'. "ACER Spatial abilities" no.1	20	8.95 (2.03)	9.23 . (1.82)	353
Post-test: "ACER Spatial Abilities'. no.2	20	10.77 (2.91)	9.32 . (3.13)	886
Post-test: "The mountains"	4	2.95 (1.13)	2.36 (1.26)	895
Post-test: "Remote Driving"	16	13.86 (3.18)	10:.36 (4.02)	.997

Notes: 1 Each group contained 22 pupils (44 in all).
2 Numbers in brackets are standard deviations of the score distributions about the given means.
3. Tabulated values of : are probabilities that the differences between the given means did NOT occur by chance.

4.4 DISCUSSION

The principal cognitive goal of the Projection Module is enhancement of projection ability. This we seek to measure by examining skill in particular spatial tasks. There are spatial tasks that are expeditiously solved using projection. However in many such tasks there are alternate algorithms not primarily based on projection. Thus in everyday life, many people with most inadequate projective ability successfully use road maps by rotating the map until the section of road they are approaching is straight ahead. In contrast, a person with good projective ability, can crawl (in his mind) along the streets on the map, irrespective of their heading.

The Remote Driving test is probably the best measure available of "pure projection". However, in this test a good score could be achieved by learning a set of "rules of thumb" such as the following: "if coming from the right (of the plan relative to the viewer) then a car turns right to go directly away".

In both the primary school and high school experiments, there was significant improvement in Remote Driving score. In the primary experiment, where Remote Driving was used as both pre-test and post-test, significant enhancement was apparent both by comparing the experimental group before and afters and by comparing experimental and control in post-test. The results also show that with the primary school children, the experimental group was marginally inferior to the control group

at remote driving up the pre-test.

The "Left-Handed Route" problem is not a test purely of projective ability, as some measure of planning skills are Involved. In the pre-test, only 4 control group pupils and 3 experimental group pupils completed the task entirely correctly; in the post-test 21 of the experimental group of 27 were successful, versus 3 of the control group (also of 27). with an increase of this magnitude, the application of statistics is indeed frivolous: one can assert with confidence that the improvement was large.

Given the success of our primary school experimental subjects on the :left-handed route: problem, we naturally anticipated similar success by the high school subjects. However, when tested, our high school experimental group performed only slightly better than the control group (10 successes versus 8). Since the experimental group had already demonstrated its greater ability at projection on the remote-driving test, this insignificant difference warrants some informal observations. In presenting this test along with others, it was not possible to ascertain whether failure to complete the task was because of lack of time or of inability. And was the inability due to lack of understanding of the written question, rather than of the abilities needed in solution?

The score distributions of the primary students in the Mountains test is presented in Figure 3. The subjects in the experimental group performed marginally better overall on the mountains puzzle than their control group counterparts. Only four particular questions were asked and there was a large spread in performance with the average score about 1 out of 4 correct in pre-testing. On the post-test, the average score was nearer 3 for both groups. The somewhat greater improvement of the experimental group was not statistically significants masked as it was by the pronounced training effect. Note our remarks above re the use of pure projection versus other means to solve the Mountains task. It is indeed likely that repetition provides opportunity for subjects to discover solution algorithms alternate to projection. In the high school experiment, where a printed version of the Mountains was used only in the post-test, the experimental group was significantly superior (at the 10% confidence level).

The ACER spatial abilities tests were a pot pourri of questions, not specifically tailored to testing projection. The improvement of experimental group over control was of statistical significance at (about) 10% confidence level. had this test not been included in the post-test, the number of questions included in the high school Mountains, Remote Driving, and Left-Handed Route could have been greatly increased, reducing the variance of the test scores.

Poor reading comprehension obviously did affect the scores on the printed tests given to the high school pupils. We suspect that lack of comprehension may have been involved in a few cases of well nigh failure amongst the primary school children too. About one third of the experimental group consisted of migrant children. While all of them were apparently fluent English speakers, improvement was generally lower than for the rest of the experimental group.

In allotting high school students into E and C groups the students were first paired. To select which student in a pair Went into the E group, it is clear (from ACER No 1 scores) that we tended to put the weaker of the pair into E. A better match of E and C

would have been achieved by tossing a coin to make that selection. A bigger sample size would have ennobled separate study of (non-English) migrants and would aid in untangling the other factors applying in Experiment 2.

In the experiment 2, only the total score in the ACER test has been subjected to analysis. However it is possible to unambiguously classify some of the questions included in the ACER tests as involving Topology? Reflection, or Rotated Figures Recognition. Other questions admit of a less ambiguous classification or else involve these skills in combination with other problem-solving factors. Such a question classification could be used to extend the above analysis of the high school experiment.

One must beware of reaching hasty conclusions as to the impact of the Projection Module on students of different age groups. The primary students involved in experiment 1 were of average ability. However our high school subjects were chosen amongst the weaker students in their classes. Nevertheless it is tempting to speculate whether the younger child at the Piagetian concrete operational stage can in general gain more from such an experience as the Projection Module.

Our use of psychometric tests in assessing cognitive change is consistent with current trends in cognition (Carroll, 1974; Resnik, 1977). Cronbach et al (1970,1974) have presented discussions of experimental design especially relevant to studies applying this strategy. In the conduct of Experiment 1, the same tests were used for both pre- and post-tests. This aspect of design was essentially superfluous, as in comparing improvement we had available experimental and control groups, but saved us from the need to devise other tests or criteria (class marks in maths etc) to select experimental and control group members. The situation re the Mountains test in Experiment 1 is well brought out by the plotted results in Fig 3. There was a pronounced training effect, which masks any improvement. The variance was also increased be the small number if questions asked. It is clear that this test is one whose reliability would have been improved by restriction to the post-test.