Integrated Learning Systems: An Analysis Using Data From Monitoring Systems

CEM Centre, Durham University

A REPORT FOR PHASE 3 OF NCET NATIONAL EVALUATION

October 1997

John Hulme, Neil Defty, C.T.Fitz-Gibbon

Some Comments Made By Students Involved In The Phase 3 Evaluation

I find it easy, exciting and easy to learn.

Year 9 Student, Institution I

I think it is fun.

Year 9 Student, Institution L

I dislike computers. I can't use them easily and they breakdown too easily.

Year 9 Student, Institution D

It is quite boring but informative.

Year 9 Student, Institution D

Computer work is fun at times.

Year 9 Student, Institution G

I found the ILS did not benefit me in maths what so ever.

Year 9 Student, Institution D

It is a little hard but it makes your brain work!

Year 9 Student, Institution F

Working on the ILS is sometimes a bit hard to understand the American language when the man talks and the money is difficult.

Year 9 Student, Institution I

List Of Abbreviations and Acronyms

Abbreviation / Acronym	Meaning
CCC	Computer Curriculum Corporation
CEM	Curriculum, Evaluation and Management Centre
GCSE	General Certificate of Secondary Education
GE	Global English
GM	Global Maths
ILS	Integrated Learning System
IPM	Initial Placement Motion
JA	Jostens Advantage
KS3	Key Stage 3
MCS	Maths Concepts and Skills
MidYIS	Middle Years Information System
NCET	National Council for Educational Technology
RI	Reading Investigations
RM	Research Machines
RW	Readers Workshop
SIR	Systems Integrated Research
SM	SuccessMaker
YELLIS	Year Eleven Information System

List of Tables In The Report

List of Figures In The Report

List of Tables In The Annex

List of Figures In The Annex

28	Graph of standardised residual in Mathematics versus total time on system
29	Graph of standardised residual in Mathematics versus total score on system
30	Graph of total time on system vs Mathematics Standardised residual
31	Graph of total score on system vs Mathematics Standardised Residual
32	Scatterplot of Time logged by MCS against Standardised Residual
33	Scatterplot of Total Achieved Score logged by MCS against Standardised Residual
34	Standardised MidYIS Test scores by use of MCS
35	Achieved Mathematics Residual by use of MCS
36	Scatterplot of Time logged by MCS against Standardised Residual
37	Scatterplot of Total Achieved Score logged by MCS against Standardised Residual
38	Standardised MidYIS Test scores by use of MCS
39	Achieved Mathematics Residual by use of MCS
40	Scatterplot of number of sessions on Readers Workshop vs English Standardised Residual
41	Standardised MidYIS Test scores by use of RW
42	Achieved English Residual by use of RW
43	Scatterplot of number of sessions on Reading Investigations vs English Standardised Residual
44	Standardised MidYIS Test scores by use of RI
45	Achieved Mathematics Residual by use of RI
46	Scatterplot of Time logged by MCS against Standardised Residual
47	Scatterplot of Total Achieved Score logged by MCS against Standardised Residual
48	Standardised MidYIS Test scores by use of MCS
49	Achieved Mathematics Residual by use of MCS
50	Standardised MidYIS Test scores by use of MCS
51	Achieved Mathematics Residual by use of MCS
52	Boxplot of performance in YELLIS test by use of MCS
52	Boxplot of standardised residual in Mathematics by use of MCS
54	Key Stage Performance
55	Residuals KS3-KS4
56	Boxplot of YELLIS Mathematics component by use of MCS
57	Standardised Residuals for Mathematics by use of MCS
58	Boxplot of YELLIS maths component score by usage of MCS
59	Boxplot Of Standardised Residuals for Mathematics by use of MCS
60	Graph of standardised residual in Mathematics versus total score on system
61	Graph of standardised residual in Mathematics versus total time on system
62	Scatterplot of time recorded on system against Standardised Residual in Maths
63	Scatterplot of time recorded on system against Standardised Residual in Science
64	Boxplot of YELLIS maths component score by usage of MCS
65	Boxplot Of Standardised Residuals for Science by use of Jostens
66	Boxplot of YELLIS maths component score by usage of MCS
67	Boxplot Of Standardised Residuals for Mathematics by use of Jostens
68	Institution G – Students Attitudes
69	Institution G – Student Comments
70	Institution F – Students Attitudes
71	Students comments – Institution F
72	Institution H – Students Attitudes
73	Students Comments – Institution H

74	Institution I – Students Attitudes
75	Students Comments – Institution I
76	Institution J – Students Attitudes
77	Students Comments – Institution J
78	Institution D – Students Attitudes
79	Students Comments – Institution D
80	Students Comments – Institution K
81	Institution L – Students Attitudes
82	Students Comments – Institution L
83	In your experience, how helpful have you found ILS systems to be in everyday class work?
84	Based on your experience, how well spent would you say was the money that bought the ILS system?
85	Was the time you spent using the system and reports better spent than if you had been teaching or preparing a traditional lesson?
86	For which pupils do you believe that Integrated Learning Systems are particularly suitable?
87	How would you rate your enthusiasm for ILS before and after making use of it in your teaching?
88	With the benefit of hindsight, if you were to teach this year again, how much more or less use of ILS would you make?
89	How useful do you find the reports in planning lessons and courses?
90	Do you find the system user friendly?
91	How would you describe the effect on discipline of using an ILS in a classroom environment?
92	If another school asked your opinion as to whether or not they should buy an ILS system, how would you advise them?
93	How much time did you spend training before you started to use the system?
94	Please rate how well you think that the system addressed your curriculum needs
95	How much better or worse do you believe those pupils who have used ILS have done than they would have done with no experience of ILS?
96	Did your pupils find the system user friendly?

Abstract

Integrated Learning Systems are in their infancy. This evaluation study, a small part of the Phase 3 NCET National Evaluation, represented a tough challenge for them. The aim was to use quietly ongoing monitoring systems, already in place in many schools, to look at the impact, if any, of the use of ILS on achievement in externally-set and externally-graded examinations, namely Key Stage 3 and GCSE. Additionally, there was the possibility of looking for effects on pupils' attitudes to computers as represented by their responses on a long questionnaire administered as part of the monitoring projects. A particularly relevant item was one concerned with their interest in working with people, machines, computers, animals, paper, etc. Would the experience of learning on an ILS package increase or decrease their interest in subsequent work involving computers?

Samples

For the year 11 samples, the YELLIS tests and questionnaires were used in six schools with a total of 214 pupils using ILS packages (the 'ILS users') and 415 pupils who were 'ILS non-users.' Their data could be set in the context of data from over 37,000 students participating in the YELLIS project nationally.

For the Year 9 samples, eight schools used the Year 9 MidYIS test and completed questionnaires as appropriate. These schools contained 416 ILS users and 599 ILS non-users. Since this was a pilot year for MidYIS, comparative data from the 14 MidYIS schools amounted to only 1,372 pupils. (For 1997/98 the number of pupils participating in MidYIS is over 34,000.)

Outcomes

In this evaluation of ILS packages, learning was not measured in terms of the gains associated with the system itself, but in terms of the apparent impact on *external examinations*. This was an ambitious hope since ILS packages may teach pupils effectively but the learning may not be transferred to other contexts or response modes. Specifically, the learning gains might be there but might not show up on the external examinations, particularly if the content was not specifically geared to the particular examinations. In other words there may well be a lack of *curriculum alignment.*

Similarly, the evaluation of possible impacts on attitudes was particularly demanding in that it relied on questions that were embedded in questionnaires that, as far as the students were concerned, were unrelated to the ILS packages. Thus *demand characteristics* (ie the pressure to report positively so as to please the people responsible for questionnaires) would be eliminated. Furthermore *Hawthorne effects* (the effect of special attention) would also be absent as the evaluation did not involve extra contacts and a sense of being evaluated. Students completed external examinations as part of their normal schooling and the questionnaires as part of routine monitoring.

Findings

Achievement: Under these challenging conditions it was difficult to find anything but an apparently negative impact on achievement. The pupils who had used ILS packages had generally lower grades on external examinations than those achieved either by similar pupils in non-ILS groups within their own school or by similar pupils in the large samples in the monitoring systems.

However, the apparent impact could have been affected by the fact that pupils using ILS packages were often dissimilar in a variety of unmeasured ways from pupils not using these packages. There appeared to have been a tendency to use ILS packages with lower achieving students and, although this could be measured and adjustments made to yield 'value added' scores, such statistical controls are always imperfect. Because of a lack of randomised assignment to ILS use or non-use, there was no equivalence between ILS users and ILS non-users. This is typical of much research into educational practice and leaves conclusions highly vulnerable to the artefact of which particular children were assigned to which particular method of instruction. (This concept is often summarised as *correlation is not causation.)*

There was some promise that the Year 9 mathematics systems could have greater impact if used more intensely since there were small but rather consistent positive correlations between the value added scores and the time students had spent using the ILS packages.

Attitudes: Pupils using ILS packages were significantly more likely to express higher levels of interest in working with computers in their future careers than pupils not assigned to ILS packages. Although it seems reasonable to attribute this finding to positive experiences of computers in the ILS packages, once again the interpretation was threatened by possible non-equivalence of ILS users and non-users. However, it seems a fairly reasonable outcome to expect from further familiarity with today's computers, whether through ILS or by other means.

Additional information, including observations and discussions, led the evaluators to the conclusion that, as ILS packages struggle through their infancy, considerable attention needs to be paid to their *mode of use*. There have been substantial investments in ILS materials, (eg £20,000 per school). It will be essential, if there is to be value for money in the long-term, to find methods of use, that are as effective as possible. This may require staff development or other ways to improve the impact of the systems. **It is strongly recommended that experiments should be undertaken to explore the effectiveness of different *methods of use* with existing systems before further investments are contemplated. Suppliers might also be required to publish results of well controlled field experiments just as pharmaceutical companies have to provide evidence of safety and effectiveness prior to marketing treatments.**

Well controlled randomised experiments of **use** versus **non-use** of ILS packages was previously undertaken by the University of Leicester (under the direction of Dr Jean Underwood) and these demonstrated learning gains from ILS materials. The positive cognitive gains found in those controlled experiments suggest that it would be worth some further investment in terms of time and effort to discover ways to make ILS packages effective enough to produce visible impacts on examination outcomes, i.e. on the broader achievements that are used in the UK to measure the effectiveness of schools and the merits of pupils.

Meanwhile it seems that Integrated Learning Systems are not guaranteed performance enhancers, nor guaranteed motivators although they do seem likely to have enhanced pupils' general interest in working with computers. They are tools that require hard work to achieve results and require further development.

1. Background

The Curriculum, Evaluation and Management Centre at Durham University participated in Phase 2 of the National ILS Project co-ordinated by NCET, producing the report "Using YELLIS to Investigate ILS Materials" which was submitted in February 1996.

The CEM Centre submitted a proposal to NCET for a small-scale research project for Phase 3. The project was to be undertaken with particular reference to the effect of ILS use on performance of students as measured by the CEM Centre's monitoring systems that provide 'value-added' reflecting students' progress towards the assessments arising from the National Curriculum key stage tests and public examinations (GCSEs).

This proposal was accepted in modified form and the project for the Phase 3 evaluation was begun in August 1996.

2. Methods

2.1 Overview of evaluation

An integrated learning system is a computer based system which delivers course material to students and tests their comprehension of this material. The delivery of new material is based on the performance and progress of the student as recorded by the system, so that each student receives an individual program of work. Teachers are able to monitor the performance of students and provide additional work and teaching.

We used existing value-added monitoring systems to investigate the effect of using Integrated Learning Systems on the learning gains of students at Year 9 and Year 11.

At Year 9 we used the newly developed Middle Years Information System (MidYIS) and at Year 11 we used the established Year Eleven Information System (YELLIS).
The two tests provided baseline measurements of general developed ability. A large background of students who had not used Integrated Learning Systems in the previous year would be used for comparison.

Table 1 – Numbers of students involved in the evaluation

	Total Number In Project	Number Of Students using ILS
YELLIS (Year 11)	37,458	214
MidYIS (Year 9)	1,372	416

We also planned to use the publicly available Key Stage Levels of the students involved as an alternative baseline. The output measure we were to use to measure performance gain was to be Key Stage 3 Levels for the Year 9 students and Key Stage 4 Levels (GCSE Grades) for the Year 11 students.

The schools that had Year 9 and Year 11 students using ILS in 1996-97 were mainly using mathematics and reading packages from three firms. These systems were SuccessMaker, Global and Jostens Advantage, and these are the systems we evaluated.
Our use of YELLIS and MidYIS also allowed for a qualitative analysis of attitudes.
The attitudinal questionnaires the students completed as part of these systems were augmented with questions more specifically directed at the use of Integrated Learning Systems by the students.

To further illuminate our quantitative findings we also conducted a survey of the attitudes of teachers towards ILS in our evaluation schools.

Initially it was hoped that we would be able to measure the effect of Integrated Learning Systems as a single intervention. As the evaluation progressed, it became clear that the use of ILS in a school was considerably more involved than it would first seem, and consequently generated far richer data. A change of approach became necessary. Specifically we decided upon the following:

- Although the three IL systems in our evaluation fell into the very broad definition of an integrated learning system given above, they were sufficiently different to be treated separately. It seemed likely that to talk of "an ILS effect" would be to over-simplify a complicated matter. We decided to analyse the systems independently of each other.

- The differences in the systems meant that there were strategic differences in implementation. Each system developed data in a different way, which meant that data sets were not easily comparable.

- Each institution had its own ideas about implementation and the use made of developed system data. There was no consistency in implementation. There were no experimental and control groups using ILS within or between institutions.
  We decided to look at each school individually as well as making careful use of groups of schools.

2.2 The Evaluation - Residuals, Effect Sizes, Indicator Systems

Residuals

Our primary unit of analysis was the student level residual. Within both MidYIS and YELLIS, each student generates a test score which can be used to predict the performance of that student at the relevant Key Stage test. When the student gains an actual grade at this level, a *student level residual* can be calculated. This is the difference between the actual achieved grade and the grade predicted by the test. These residuals allow for fair comparisons between subjects, schools and users or non-users of IL systems.

Effect Sizes

An effect size is the standardised difference between the means of the experimental and control populations. In the present study, the 'control' or comparison group was taken to be those students who had completed MidYIS and YELLIS tests and who had almost certainly not used Integrated Learning Systems in their education in the school year 1996-97. Effect sizes are a useful tool of analysis for three main reasons:

- They allow for comparisons between the impacts of unrelated interventions.
- They are easy to understand.
- They provide a quantitative measure of effect that is not based on an arbitrary notion of statistical significance.

Meta-analysis is particularly suitable for this kind of comparison, as Fitz-Gibbon (1984) comments " (one essential feature of passive observational studies is) the examination of data for relationships which *suggest hypotheses*. They are distinguished from the activity of *testing hypotheses* for which one needs, ideally, controlled experiments."

Of further use to us was the fact that effect sizes had previously been used as a tool for the evaluation of Integrated Learning Systems (Underwood *et al.* 1994), and this would allow our results to be quickly placed into context with previous findings.

Indicator Systems – Providing a background

Our use of established and efficient indicator systems meant that we could call upon a large statistical 'background' to our evaluation. This supplied the control groups for the effect sizes analysis and also provided questionnaire responses for our qualitative study of attitudes.

2.3 The YELLIS test

The YELLIS project has two modes of participation, the two-year cycle and the one-year cycle. In the two-year cycle, the YELLIS test is taken in the first term of Year 10. In the one-year cycle, the YELLIS test is taken in the second term of Year 11. The YELLIS test is composed of a vocabulary test, a mathematics test, and a short attitudinal section. Participants in the two year cycle have the option of taking the Extended YELLIS questionnaire which contains extensive attitudinal questions.

All but one of the schools in the Durham component of the Phase 3 evaluation were part of the one year cycle. The other school was already involved in the YELLIS project, and its Year 11 students had completed the YELLIS test at the start of Year 10. Appropriate procedures were applied to the analysis of data from this school.

All the students in the ILS group took the basic YELLIS test. Where it was possible, the basic test was supplemented with further attitudinal questions which were a subset of the Extended YELLIS questionnaire. The supplemental attitudinal questions dealt with attitudes to computers, machinery, various school subjects and careers etc.

2.4 The MidYIS test

The MidYIS and YELLIS tests are similar in gross structure. Both tests have a section that measures general attained academic and cognitive ability, and an attitudinal section.
The MidYIS test has 5 'academic' components that fall into three broad categories vocabulary, mathematics and non-verbal skills. MidYIS test students were also given an attitudinal questionnaire based on the basic Yellis questionnaire and the subset of the Extended questionnaire which had been made available to the Year 11 students.

Like the YELLIS test, the MidYIS test is capable of being used as a concurrent baseline test, because its content is largely curriculum independent.

2.5 Additional Sources of Information

Several other sources of information were available to us. We administered a questionnaire to staff who used ILS in their teaching which provided rich qualitative data about methods of implementation and attitudes to use of the systems.

A database of information about the evaluation schools was made available to us by NCET. It was continuously updated and was another extremely useful resource for information about modes of use and for providing qualitative data about the institutions.

NCET also provided summaries of reports made after visits of the schools by their project officers, which proved to be another very useful source of qualitative data.

2.6 Timetable of Evaluation

We referred to the schools we were to evaluate using YELLIS and MidYIS as the Durham Distributed Research Group (DDRG). The schedule of events was as follows:

1996	
September – December	• Identification of Durham Distributed Research Group from NCET database • Schools in the DDRG contacted • Control sample questionnaires distributed

1997	
January - March	• YELLIS testing • Control sample identified • Staff attitudes questionnaire distributed
April - June	• MidYIS testing • Student Attitude Questionnaires distributed
July – September October - November	• Collection of System Data and Results data • Preparation of Final Report • Synoptering

2.7 Considerations and Complications

The analysis environment was both diverse and dynamic. Things rarely worked in practice as they were described on paper. This caused us to change some of our analysis strategies, and also gave rise to an awareness of several important factors that needed to be considered when interpreting our analysis. These important considerations were:

- Nature of the Intervention
- Differences in curriculum alignment
- Differences in implementation of ILS
- Differences in the nature of the recorded data
- Lack of data
- Analysis overload and the co-operation of schools

Nature of the Intervention

As a properly controlled experiment, the evaluation could have looked solely at ILS as a single intervention but what we have actually evaluated is a group of several different interventions. These interventions included the teaching practices of the school, the nature of the school itself and the teachers who taught the students in the evaluation as well as the Integrated Learning System and the methodology of its implementation.

In other words, Integrated Learning Systems were only part of the package of educational provision which these students received and effects noted between groups of students who did and did not use ILS should be interpreted in this light.

Differences in Curriculum Alignment

The extent to which the content of the material delivered by the ILS coincided with the content of the external end-of-key stage examinations varied between IL systems. Since this factor alone could be expected to account for considerable differences in the apparent impact of the external examination, it was important to consider the extent of curriculum alignment.

Differences in Implementation of ILS

- There was no consistency in the enrolment of students. For example, some schools enrolled a whole year group onto a system, whilst others chose to enrol just the lower sets. e.g. Institution D, Institution C. No schools chose to enrol students onto the system as part of an experimental group for which there was a control.

- The timetabling of ILS use was variable. For example 'Open Access' implementations meant that data was either not recorded or of very poor quality. Some schools wrote use of the ILS into the timetable, whilst others allowed students to use the ILS when required as a revision aid.

- HMI said of ILS "ILS is an approach which critically depends on computer feedback based on performance and the management of the curriculum of the individual student." SuccessMaker, Global and Jostens all have very different approaches to feedback. There are qualitative and quantitative differences in the data generated by each IL system. SuccessMaker generates the largest volume of data in terms of information. It measures and reports many different aspects of a student's use of the system, including the total number of sessions completed and the percentage of correct answers achieved as well as a breakdown of all the different elements which go to make up a SuccessMaker module. Jostens generates similar data to SuccessMaker, but in much less detail. This is largely because Jostens modules do not share the 'strand' nature of SuccessMaker modules. Global Mathematics produces the least reporting data in terms of information (although the most in terms of physical bulk) which details only the time a student spent on each module, and the final score achieved in the module.

- There were varying levels of support for ILS within an institution. Sometimes enthusiasm for ILS was part of the culture of the school, in others it seemed to be stoutly defended by just a few teachers.

Differences in the nature of the recorded data.

- Each system records data about usage and performance differently. One critical difference is 'time on system' For example, Global records just the time the student spends completing tasks whereas SuccessMaker records the time the student is actually logged onto the system. This time is different again from 'time on task'. The breakdown of performance data is also very different, Global reports the achieved percentage in each module, whereas SuccessMaker has a very diverse reporting system which can format large quantities of information about all aspects of a student's performance on the system. The relative merits of these approaches to reporting are not immediately obvious.

- Reinstallation of systems means that whilst nearly all students spent a full school year on the system, several schools were only able to provide a few months worth of data (Institution F Year9, Institution B).

- The systems recorded a student accessing a module, whether or not any work was completed, or indeed attempted. This means that a certain percentage of system data will inevitably have been generated by students who were 'just having a look' rather than who were unable to complete the task.

Lack Of Data

There were four major factors which contributed to a lack of data:

- Five sets of data were associated with each ILS student. These were Key Stage Input, YELLIS/MidYIS test details, Key Stage Ouput, System Data and Attitudinal Data. Problems with the recording and availability of data meant that not every student had five complete sets of data. This means that the numbers of 'valid cases', or students suitable for analysis vary between analyses as cases were deleted pairwise rather than listwise. The number of valid cases is presented with each analysis.

- Global English was not able to export or print out reports of system usage. For this reason, the time on system data for Global English has been extrapolated from timetable information provided by the schools involved.

- Key Stage 2 information had not been recorded, or it had not followed students to their secondary schools. This meant we were unable to use Key Stage 2 as an alternative baseline to the MidYIS test for the Year 9 students.

- Data was lost from the evaluation due to system crashes and accidents during upgrading of some systems. Examples include an institution in which the system hard drives were accidentally reformatted whilst upgrading the operating system and Institution B that had a system crash mid-year and had to re-enrol students in April. Several other institutions had such severe problems with hardware and software that they were unable to be included in this evaluation. This group includes the institution whose system was 'hacked' by a student. The co-ordinator wrote "... the SuccessMaker system is not a secure system when running under

Windows 95 ... One of the students ... went to the history disk and deleted all the data ... (by the time I had noticed) we had gone through our weekly set of backup tapes and were backing up empty areas."

Analysis Overload and the Co-operation of Schools

- There were three evaluation teams. In addition to the CEM Centre, Durham, another evaluation entailed teacher-researcher projects and a third involved large scale administration of CAT tests, each taking two and a half hours, in addition to the recommended preparation periods. Some schools were involved in all three aspects of the evaluation. The workload on these schools was particularly heavy. The heavy workload and the fact that three independent evaluations were taking place meant that some co-ordinators were confused as to what they should be doing, when, and for whom. The work imposed on the co-ordinators and the institutions themselves, meant that they were sometimes less able to be as helpful as they might otherwise have been.

3. Data Sets

Data sets were located from the NCET database using information which had been supplied by the schools themselves at the initial conference.

3.1 Year 9

System Data There were two major systems represented, SuccessMaker – manufactured by the Computer Curriculum Corporation, and Global from Systems Integrated Research.

The data that was returned to Durham for SuccessMaker systems was mainly for the Maths Concepts and Skills module, although very small amounts of data were returned for the Readers Workshop and Reading Investigations modules.

Global data returns were for Global Mathematics. Some data was obtained for students using Global English, but the system data itself was irretrievable.

Table 2 – Summary of Year 9 ILS System Data Set

Provider Of ILS	Name Of ILS	ILS module	No. Of Schools	No. of Students	Curriculum Area
SIR	Global Mathematics	-	2	179	Maths
SIR	Global English	-	2	28	English
CCC	SuccessMaker	Maths Concepts and Skills	4	176	Maths
CCC	SuccessMaker	Reading Investigations	1	31	English
CCC	SuccessMaker	Readers Workshop	2	63	English

MidYIS Test and Attitudinal Data -All Students

All the students in the institutions involved were supplied with Year 9 MidYIS tests and additional questionnaires. The following table summarises the data set which was returned.

Table 3 – Summary of Year 9 Questionnaire Data Set

School	MidYIS	Additional Questionnaires	% of MidYIS sample completing questionnaire
Institution G	151	150	99
Institution F	116	100	86
Institution H	157	137	87
Institution I	79	45	57
Institution J	78	72	92
Institution D	171	168	98
Institution K	146	46	32
Institution L	10	11	100

3.1.3 Key stage 3 data for Year 9 Class of 1997

There were some problems with the collection of this data. There seemed to be widespread dissatisfaction with the English KS3 results which meant that most schools had their English results re-marked which resulted in delay.

Table 4 – Summary of Year 9 Students with Input and Ouput Measures

Category	Students
Number of ILS using Students with KS3 output	357
Number of ILS using Students with MidYIS data	283

3.2 Year 11

System Data

There were two major systems represented for Year 11 students: CCC's SuccessMaker and SIR's Global. There was also one school that used Jostens Advantage.

The data that was returned to Durham for SM systems was for the Maths Concepts and Skills, Readers Workshop and Reading Investigations SuccessMaker modules. Global returns were for Global Maths.

Table 5 - Summary of Year 11 ILS System Data Set

Provider Of ILS	Name of ILS	ILS module	No. of Schools	No. of Students	Curriculum Area
SIR	Global Mathematics	-	2	127	Maths
CCC	SuccessMaker	Maths Concepts and Skills	3	69	Maths
CCC	SuccessMaker	Readers Workshop	1	12	English
Jostens	Jostens Advantage	Maths and Science	1	20	Maths and Science

YELLIS Test and attitudinal data sets – All Pupils

All the students in the institutions involved were supplied with Year 11 YELLIS tests and additional questionnaires. The following table summarises the data set which was returned.

Table 6 – Summary of Year 11 Test Data Set

School	YELLIS
Institution A	86
Institution B	169
Institution C	96
Institution D	175
Institution E	25
Institution F	78

Key stage 4 (GCSE) Data for year 11 class of 1997

GCSE results were requested from all those who had taken YELLIS tests. The following table summarises the numbers of students using Integrated Learning Systems at Year 11 for whom input - and output data was returned.

Table 7 – Summary of Input and Output Data for Year 11

Category	No. Students
Number of Students using ILS KS4 ouput	227
Number of Students using ILS with YELLIS data	227

Key Stage 3 Data for Year 11 class of 1997

We had anticipated some difficulty in collecting this as Key stage tests have only been widely or universally taken in the last few years and results have been imperfectly recorded, however unlike Key Stage 2 data, KS3 data was available for most students in the ILS schools.

Table 8 – Summary of available Key Stage Data for Year 11

School	Maths KS3	English KS3	Science KS3
Institution A	70	69	69
Institution B	0	0	0
Institution C	24	9	23
Institution D	171	0	0
Institution E	0	0	0
Institution F	52	53	56

3.3 Backgrounds for comparative analysis

YELLIS

The dataset available for comparisons outside the ILS schools was derived from the 37,358 Year 11 students who took the YELLIS test in 1997 and the 21,004 students who completed the questionnaires for the extended attitudinal background.

The YELLIS regression analysis was rerun, with those schools that had identified themselves as using Integrated Learning Systems removed from the analysis. The effect of this removal was negligible. Where possible and appropriate, the YELLIS analysis has been repeated using regression equations from the 1996 Value Added National Project. These regression equations predict KS4 (GCSE) performance from KS3 performance. The number of students who were used to generate these equations in the Value Added National Project was 25,843. (Vincent, 1997; Fitz-Gibbon, 1997)

MidYIS

The background dataset for year 9 students was derived from the 1,372 Year 9 students who took the MidYIS test. No schools in the survey of the Year 9 background had identified themselves as using Integrated Learning systems, so the whole MidYIS sample was used as a background.

4. Analysis Of Data

4.1 Summary of Effect Sizes

The following tables detail the achieved residual for each school, in the subject or subjects which were the subject of the ILS use. If a residual is negative, this implies that the group of students has not done as well as expected *all other things being equal*. If the students made the sort of progress that could be expected then the achieved residual will be zero. If the students made greater learning gains than could have been expected from their performance in the YELLIS or MidYIS tests, then they will have positive residuals.

Table 9 – Summary of Effect Sizes by Institution: Year 9

Institution	System	Subject	Apparent advantage to ILS?	Pupils Using ILS		Pupils Not Using ILS	
				No.	Mean SR	No.	Mean SR
I	RW/RI	English	Yes	68	-0.10	5	-0.37
L	GE	English	N/A	1	-4.61	N/A	N/A
J	GE	English	N/A	N/A	N/A	58	-0.05
I	MCS	Maths	No	46	0.17	30	0.33
D	MCS	Maths	No	36	0.21	130	0.53
G	GM	Maths	=	121	-0.32	23	-0.28
K	MCS	Maths	No	7	-0.33	124	0.08
H	MCS	Maths	No	21	-0.44	97	-0.17
F	GM	Maths	No	19	-0.80	93	-0.18

SR – Standardised Residual

The column labelled 'Apparent advantage to ILS' indicates the comparison of the residuals within the school. Only for Institution I's English was the residual substantially better for the ILS users than the non-ILS users in the school, and even there it was simply less negative.

Table 10 - Summary of Effect Sizes by Institution: Year 11

Institution	System	Subject	Apparent advantage to ILS?	Pupils Using ILS		Pupils Not Using ILS	
				No.	Mean SR	No.	Mean SR
F	GM	Maths	No	40	-0.12	29	0.17
B	JA	Maths	=	32	0.41	136	0.44
A	GM	Maths	No	81	-0.43	3	-0.25
D	MCS	Maths	No	33	-0.71	138	0.58
E	MCS	Maths	No	4	-0.83	20	-0.45
C	MCS	Maths	No	18	-1.09	59	-0.13
B	JA	Science	=	33	0.13	136	0.07

SR – Standardised Residual

Further representations of the data, school by school, are provided in Annex A.

4.2 Time On System

Time on system is a more involved concept than it might at first seem. We identified the following measures of time spent using the ILS.

- Time on task on the system
- Time logged onto the tasks on the system, rather than the learning components
- Time logged onto the system, but not necessarily working or learning

Time on task could only properly be established by actually observing classes of students at work, but fieldwork on this scale was outside the scope of our evaluation. This meant that our main indicator of time on system, was the time recorded by the ILS itself but, unfortunately, there were important differences in this information between systems. SuccessMaker recorded the time that students were logged onto the system. If a student spent a forty minute lesson period using SuccessMaker to learn and do exercises, then the system would record forty minutes of use.
Global Mathematics recorded the time that students spent doing tasks. If a student spent twenty minutes being taught a concept, and then spent five minutes completing a task based on that concept, the system would record a time of five minutes.

Year 9

We looked at the correlation between the time spent on the system and the performance of students in terms of their KS3 mathematics results. Using the full year 9 MidYIS sample there was a positive correlation of almost 0.3, (as strong as the usual correlation between achievement and socio-economic status).

Figure 1 - Global Maths Standardised Time On System Against Standardised Residual For Mathematics

N = 140
r = 0.27

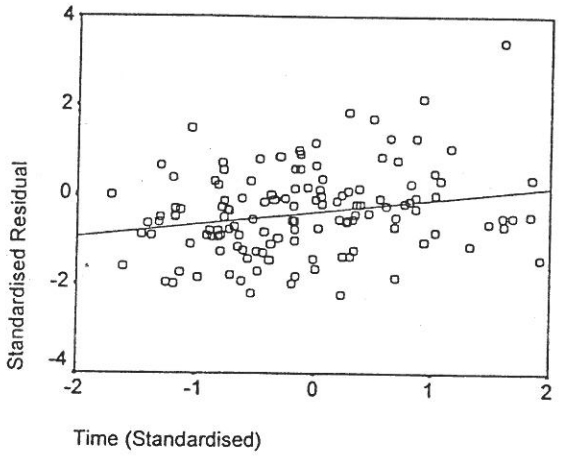
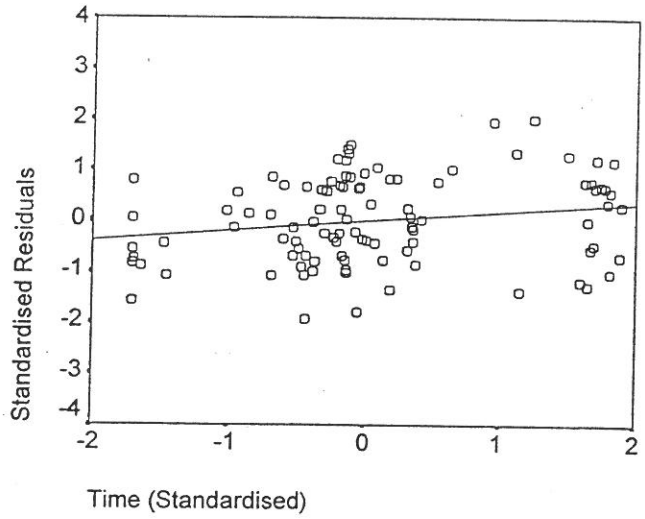


Figure 2 - SuccessMaker MCS Standardised Time Against Standardised Residual For Mathematics

N = 108
r = 0.23

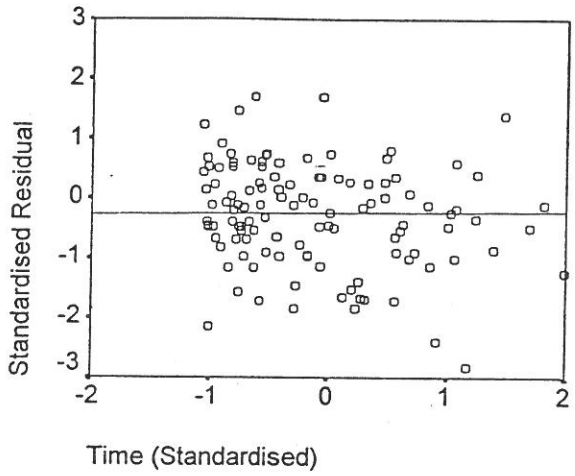


Year 11

The following graphs detail the performance of students in terms of their achieved standardised residuals on GCSE examinations compared with the *standardised* total time spent on system. This measure is, perhaps, the best measure of the relative effectiveness and efficiency of the two systems for which we had large numbers of students in Year 11 – Global Mathematics and SuccessMaker MCS.

Figure 3 - Global Maths Standardised Time Against Standardised Residual For Mathematics

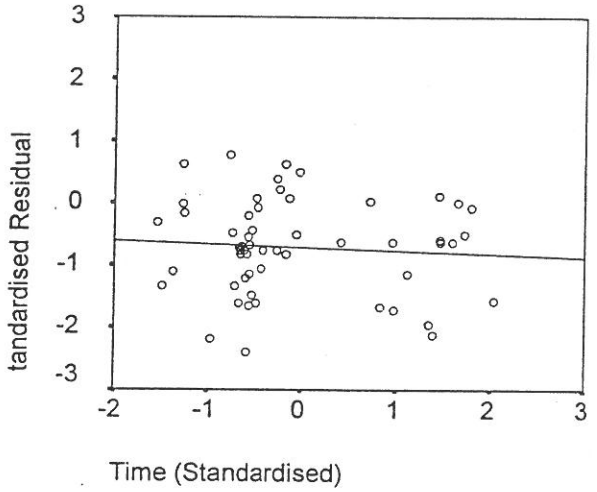
N = 121
r = 0.01



The plot for Global Mathematics indicates essentially no relationship between the time spent on the system and the achieved standardised residual at this level. The correlation between time on system and achieved standardised residual is 0.01, effectively zero.

Figure 4 - SuccessMaker MCS Standardised Time Against Standardised Residual For Mathematics

N = 55
r = -0.07



The plot for SuccessMaker Maths Concepts and Skills also shows that there is essentially no correlation (-0.071) between the time spent on the system and the achieved standardised residual.

4.3 Performance On System

The measure of performance used here is the total number of questions completed correctly, or the total score. This is in effect a function of time. It provides more illumination however, in that the system has no way of recording time on task. The use of total score allows us to measure the effectiveness of the system, when the system has actually been used by the student, and not simply contemplated for a period of time.

Year 9

We examined the actual recorded performance on the system to see if it correlated with learning gains made by the student.

Figure 5 - Global Maths Total Standardised Score on System against Standardised Residual For Mathematics

N = 140
r = 0.331

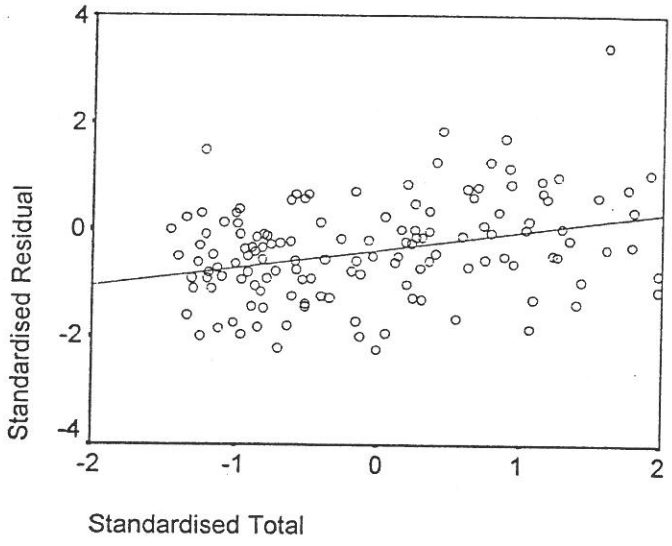
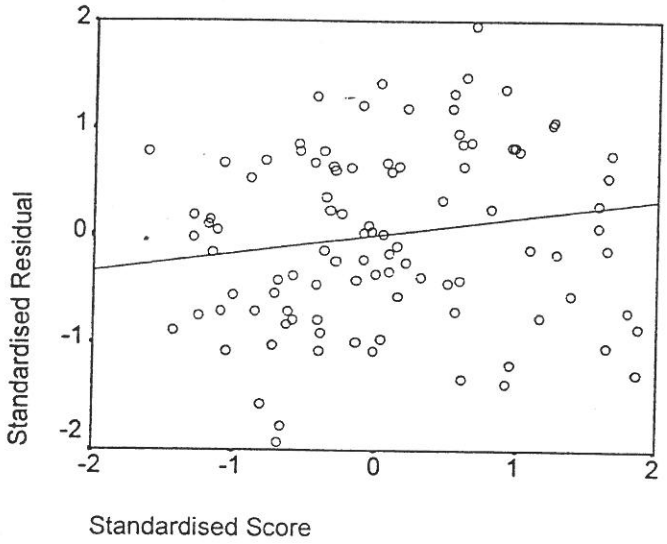


Figure 6 - SuccessMaker Total Standardised Score on System against Standardised Residual For Mathematics

N= 108
r = 0.204



Year 11

Figure 7 - Global Total Standardised Score on System against Standardised Residual For Mathematics

N = 121
r = -0.07

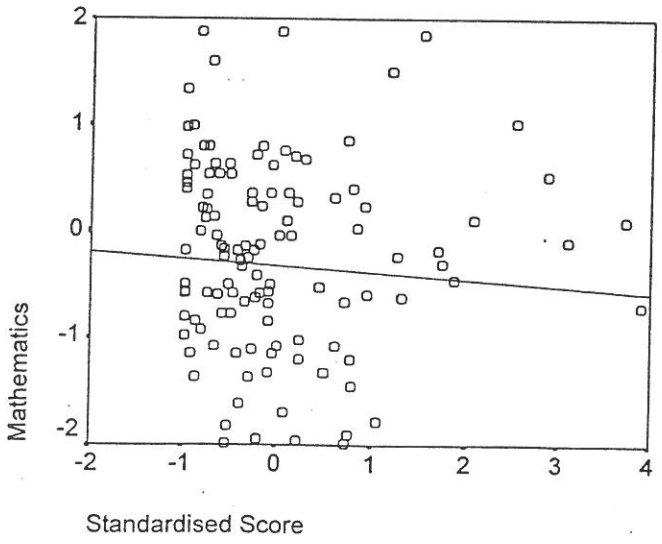
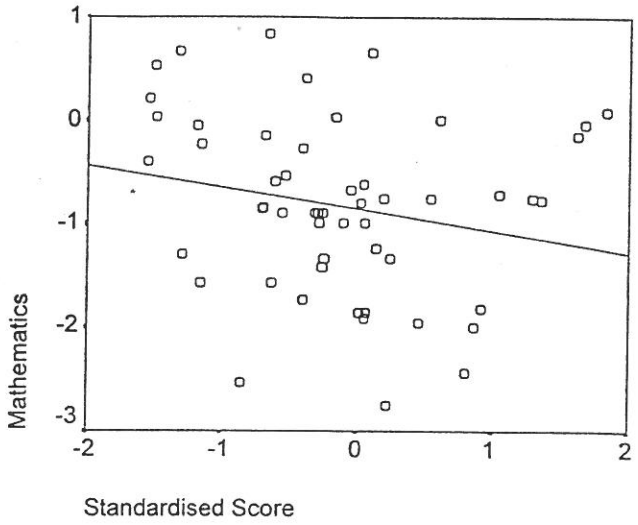


Figure 8 - SuccessMaker Total Standardised Score on System against Standardised Residual For Mathematics

N = 55
r =-0.23



4.4 Discussion and observations regarding achievement gains

For **Year 9 SuccessMaker** there were weak positive correlations between the time on task and achieved residual. For **Year 9 Global** there were weak positive correlations between time on task and residual, and total score and residual. Although objectively weak, the correlations were equivalent to figures quoted for influence of home background, which is traditionally considered to be important.

For **Year 11 SuccessMaker (Maths Concepts and Skills)** all three institutions in this part of the evaluation used the ILS system with their less mathematically able students. The actual performance of these students was noticeably lower than expected. There are several reasons why this could have happened:

- **A Confound.** It is important to remember that the effects noted are for the whole package of interventions which these students received, not solely their use of the Integrated Learning System. In other words, by using *intact classes*, often composed of the less mathematically able, any comparisons made are not entirely interpretable. The entire experience of the particular classes involved might have been different from that of other classes. In a controlled experiment, intact classes are not used unless they become the unit of randomisation.

- **Failure of Transfer.** It is possible that students learn mathematical concepts and make progress using MCS, without significantly enhancing their examination prospects. Good performance at GCSE will depend on exam technique and *familiarity* with the kind of questions asked. A particular mathematical concept, which is well understood within the framework of the ILS, might not even be recognised when recast into a different format for an examination question.
  "Some children were apparently able to do tasks on SuccessMaker which they were unable to do when presented as pencil and paper tasks after the session" Ann Lewis

- **Curriculum Alignment.** If the traditionally taught syllabus which forms the bulk of a students mathematical education and the material delivered by the ILS are not very similar in content, it may be that the time spent by the student using the ILS is less well spent than time in a traditional environment. It is important to remember that what we have measured is the students performance in GCSE Mathematics, something which depends critically on precise knowledge of the material which is to be examined. Facility in branches of mathematics which are not examined will not appear in our data set, and will not be reflected in the qualifications awarded by the examining boards. A teacher at Institution D, one of the institutions with a disparity in performance between those who do and do not use SM MCS comments "*It is very helpful but can cause some disruption to the normal maths curriculum.*"

- **Differences in response modes.** If a student is used to interacting with a multimedia computer to do mathematics, the transition to a silent examination room with nothing but paper and pencil may well have a negative effect on performance. The immediate marking right or wrong of questions by an ILS is significantly different to a blank sheet of paper and a head full of doubts. This may be particularly apparent in multi-stage examination questions where correct answers to the first stages will make the later stages easier.

- **ILS as a panacea**. Perhaps the teaching staff felt that the work required to gain good grades for the ILS students had been done by placing them into the ILS, and they turned their attention to the other students. There might well have been a particular danger of this scenario occurring if the progress of the students was measured by the SuccessMaker measure of Gain (see section 4.x.x).

Year 11 global

No Year 11 Global schools had split samples to allow for intra-school comparisons. In general it seemed that there was no correlation between time spent on the system, or performance on the system and achieved residuals at Year 11

Use of Reports: variations between systems

Responses to the attitudinal questionnaire for teachers revealed that staff who used SuccessMaker were far more likely to make use of the system reports than staff who used Global or Jostens. At least part of the reason for this may be the volume and richness of the data generated by SuccessMaker.

Our analysis of the system data and effect sizes indicated that certain elements of the SuccessMaker reporting system should be treated with extreme care. In particular, the concept of gain may well be a misleading indicator of performance. Our concern was further increased by an internal report from Institution D which planned to expand the use of SuccessMaker in the same mode of usage, based on the gains reported *by the system.*

Our analysis of the performance of the pupils who had used the system, however, indicated that, in fact, the package of interventions to which they had been subject and of which SuccessMaker had been a part, had not been successful in raising achievement in the external examinations.

"... the gain in SuccessMaker is not assessed by the administration of two identical or similar tests over an interval of time. The nature of the progression in SuccessMaker is that everyone gains (success is built in) and the gains made are simply stages along the progression."

John Gardner in Underwood, J. and Brown, J (1996)
"Integrated Learning Systems: Potential into Practice".

Figure 9 - Scatterplot of Gain as reported by SuccessMaker versus achieved Standardised Residual for Year 9 group

N= 108
r= 0.23

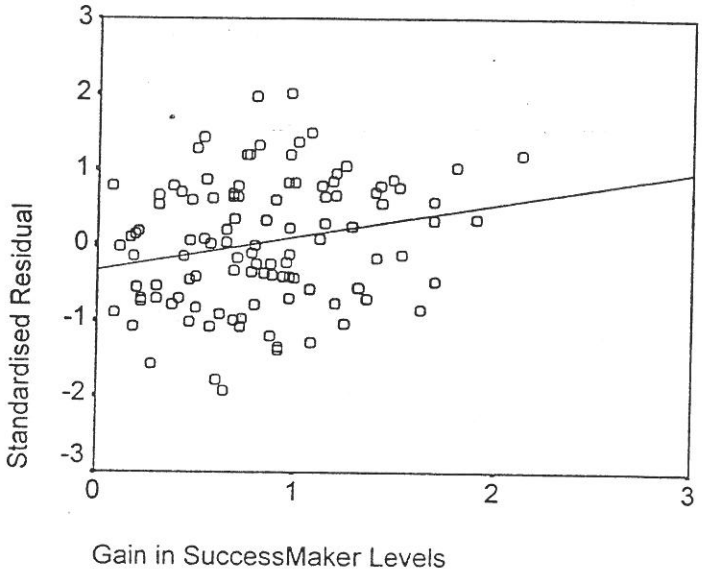
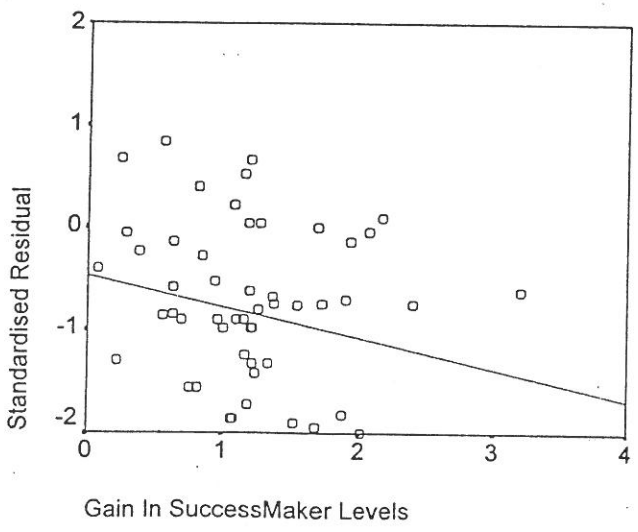


Figure 10 - Scatterplot of Gain as reported by SuccessMaker versus achieved Standardised Residual for Year 11 group

N = 55
r = -0.23



5. Attitudes as indicated by student questionnaires

The attitude questionnaires were administered following the usual procedures in the YELLIS project of providing plastic envelopes for pupils to ensure that the responses were confidential. Attitudes for pupils from the institutions involved were analysed for Year 11 and Year 9 and are presented in Annex A using the 'comparison graphs' familiar to schools in YELLIS. On these graphs the school's data is represented as bar charts and the entire data is graphed as shaded background. The background sample for questions from the Basic questionnaire in 1997 was 58,462 pupils. The background sample for questions from the Extended questionnaire was 21,004 pupils. The numbers given beneath the graphs are those for pupils with recorded ILS use.

The Year 9 attitudinal survey was a composite questionnaire derived from the basic YELLIS test questionnaire and relevant questions from the extended YELLIS questionnaire.

Year 11 attitudinal data came mainly from the basic YELLIS test. Some schools completed an additional attitudinal questionnaire which was composed of relevant questions from the Extended YELLIS questionnaire.

5.1 General Findings from Analysis of Attitudes

One question particularly relevant to students' developing attitudes towards computers asked students to indicate to what extent they would like to work, in the future, with people, animals, machines, paper, books and computers. Against the background of all the students in YELLIS and MidYIS, the responses to 'computers' were, overall, particularly outstanding. The graphs showing these responses school by school are in Annex B and one school's graph (School F, that had students in both the Year 11 and Year 9 samples) is reproduced below.

Figure 11 - Percentage of Institution F pupils expressing a preference for working with ...

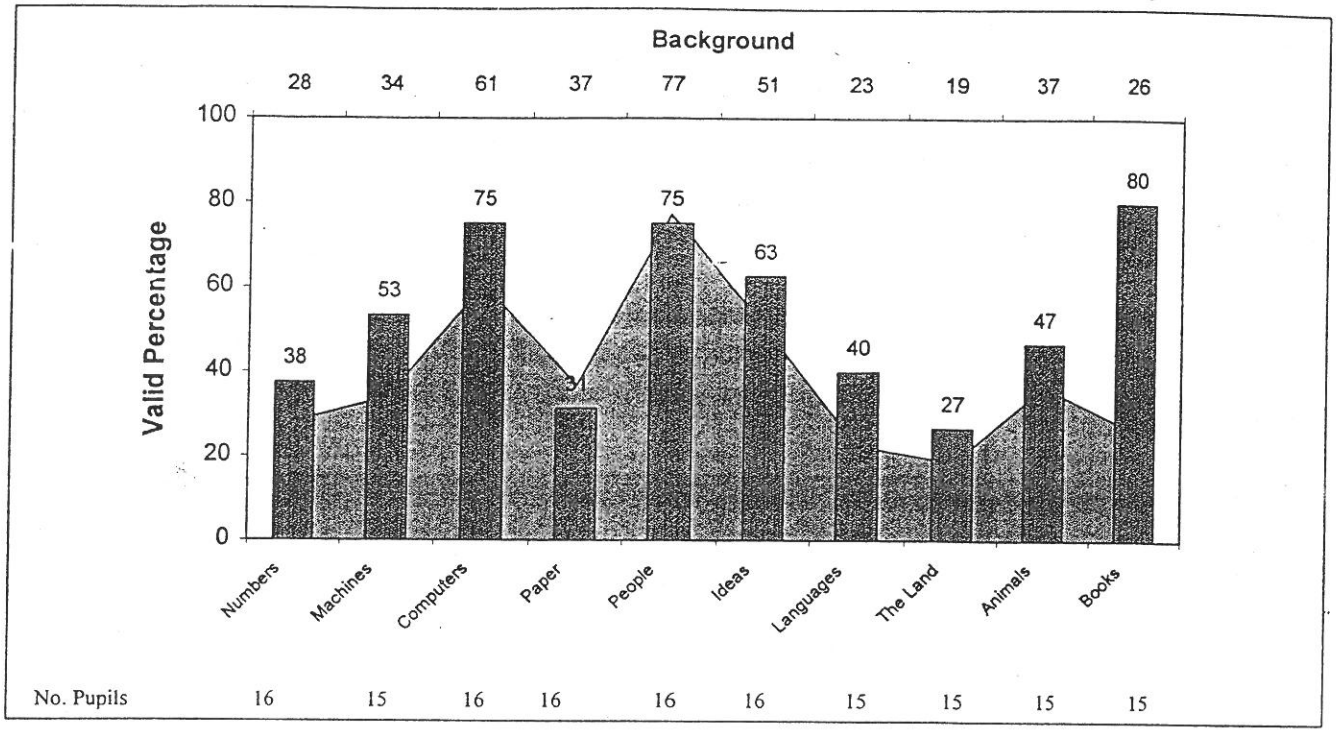


Figure 11 suggested that the percentage of pupils interested in working with computers was about 15 percent higher (75 minus 60) among the students who had been exposed to the ILS packages. Oddly, at this school, interest in books had also apparently shot up. This was not observed in other schools (see Annex B).

Averages can hide the variations in the data so a check on the statistical significance for the entire sample was made. Tables 11 and 12 indicated that the more positive attitudes to computers seemed to be associated with Global mathematics rather than SuccessMaker. The 'effects' of the CCC SuccessMaker Mathematics Concepts and Skills package on students' interest in working with computers were far from being statistically significant whereas the 'effects' of Global mathematics were significant at the .02 level. Although clearly different, the interpretations must remain tentative given the possibly numerous differences between the groups. To overcome some of the differences between groups the extent to which pupils reported highly negative views of lessons was used as a covariate (named 'alienation'). Alienation was expressed by strongly agreeing that lessons were dull and there was no point to them.

Table 11 ANOVA SUCCESSMAKER: Maths Concepts and Skills: relationship to student's expression of interest in working with computers in the future

		Sum of Squares	df	Mean Square	F	Sig.
Covariates	ALIENATION	7.967	1	7.967	6.234	**.013**
Main Effects	Use Of MCS	.293	1	.293	.229	.632
Total		907.737	706	1.286		

Table 12 The apparent effect of Global Mathematics on attitudes to computers

		Sum of Squares	df	Mean Square	F	Sig.
Covariates	Alienation	5.527	1	5.527	4.358	.037
Main Effects	Use Of GM	7.007	1	7.007	5.524	.019
Total		907.737	706	1.286		

The extent of students' reported interest in working with computers in the future was also significantly related to sex. Boys were more keen to work with computers as were those students who were least alienated from school (Table 13).

Table 13 ANOVA Interest in working with computers

		Sum of Squares	df	Mean Square	F	Sig.
Covariates	Sex	5.620	1	5.620	4.443	.035
	Alienation	6.976	1	6.976	5.515	.019
Main Effects	Global Maths	6.022	1	6.022	4.760	.029
Total		842.23	652	1.292		

In Annex C, the schools' attitude datasets are presented graphically and comments about work with computers are also included.

Students' comments about computers

The following table presents a selection of comments from the School F questionnaire supplied to year 9 pupils. The comments are reported verbatim, as written down by the pupils. No spelling or grammatical corrections have been made.

The comments are ranked by score in the MidYIS test, with the highest scoring pupil first. Each of these pupils used the Global Mathematics ILS in year 9. Pupils were asked, *"Use this box for any comments that you wish to make about your work on computers ... ".*

Table 14 Students' verbatim comments: School F, year 9.

Sex	View	Comment
F	+	I am quite good with computers. I have one at home and I use it quite often to do home work to make it neat. I use the encyclopaedia to find information.
F	+	They should give us more time on computers.
F	+	I really like working on computers, especially with CD-roms.
F	+	It is a little hard but it makes your brain work.
F	0	I think working on the computers is fun but sometimes I get a little lost.
F	+	I think that my work on the computers are that my work is neat and it is fun.

In the second column the comments have been classified as generally positive to computer use, including the expression of enjoyment or learning ('+'), mixed or neutral (0) or negative (-). Further tables of comments are presented by school in Annex B. Overall, the comments were classified as

- 52 percent positive (+)
- 22 percent mixed or neutral
- 26 percent negative (-)

Probably the most common feature of the negative comments was 'boring' although being confused was also evident, particularly towards the lower end of the tables providing comments from pupils whose test scores were lower.

Teachers' Responses to questionnaire items

A questionnaire was answered by 35 teachers. Below are the answers to a single question. A full set of responses is in Annex D.

Table 14 - Teachers thoughts on the single greatest benefit of ILS

System	In your opinion, what is the single greatest benefit of ILS?
GM	Motivational factor.
GM	Pupil motivation.
GM	Variety.
GM	The tremendous potential it could have.
GM	Individual attention to one pupil. The most closely focussed differentiation.
GM	Added motivation.
GM	Less teacher input.
GM	Pupils enjoy the interactive aspects.
GM	It motivates lower ability pupils.
GM	Reinforcement of concepts.
GM	Work in a logical progressive way on basic literacy skills to supplement traditional teaching methods - reinforcing and motivating.
GM	Reinforcement and practise of basic literacy skills.
GM	Structured individual learning.
GM	It has enabled a difficult low-ability group to be split between discipline/motivation.
JA	different learning style.
JA	Illustrates concepts which are difficult for pupils to understand.
JA	Animated programmes.
SM	Introduce motivation towards IT.
SM	Individual educational programs - S/M strands-based courses.
SM	Pupil takes responsibility for learning.
SM	Improves power of concentration over longer period of time in less able pupils.
SM	It is supportive of pupils with learning difficulties and gets them to concentrate. They learn at their own rate.
SM	Can be effective with certain pupils.
SM	It is an excellent was of teaching and providing practice particularly numeracy.
SM	The increase in motivation of my bottom set year 9 pupils.
SM	Re-enforcement and constant practices.
SM	Gives pupils individual practice on identified weeknesses in through a medium they can both understand and find stimulating.
SM	Improving basic skills.
SM	Evaluation done by computers.
SM	Increased pupil motivation!
SM	Individual motivation gained through positive statements/teacher remarks.
SM & GM	Individual progress (constantly monitored)
SM & GM	Individual needs of every student identified and catered for, and reported on.