EFFECT SIZES ASSOCIATED WITH PROLOG-BASED LEARNING ENVIRONMENTS

George P.L. Teh
&
Barry J. Fraser

Paper presented at the Sixth IFIP World Conference on Computers in Education, held in Birmingham, UK on 23-28 Jul 1995
Effect Sizes Associated with Prolog-Based Learning Environments

George P.L. Teh
Nanyang Technological University
469 Bukit Timah Road, Singapore 1025
Tel: 65 460 5180
Fax: 65 469 8433
E-mail: TEHPL@NIEVAX.NIE.AC.SG

Barry J. Fraser
Curtin University of Technology
GPO Box U 1987
Perth 6001, Australia
Tel: 619 351 7896
Fax: 619 351 2503
E-mail: ifraserb@info.curtin.edu.au

Theme : Artificial Intelligence
Level : Secondary Education
Topic : Humanities/Social Sciences
Secondary Keywords : Classroom Practice,
Computer Assisted Instruction,
Courseware, Research

Paper presented at the Sixth IFIP
World Conference on Computers in Education,
Birmingham, U.K., July 1995
23/7/1995 - 28/7/1995
ABSTRACT

This study of the efficacy of computer-assisted learning (CAL) is distinctive in that (1) the innovation in CAL made use of micro-Prolog, (2) traditional achievement and attitude outcomes were complemented by measures of classroom psychosocial environment, and (3) the sample was drawn from the unique educational milieu provided by Singaporean schools. In contrast to past research, the use of micro-Prolog-based CAL in this study led to massive impact in terms of achievement (effect size of 3.5 standard deviations), attitudes (1.4 standard deviations) and classroom environment (ranging from 0.1 to 1.9 standard deviations). These differences favoured the CAL group. Students in the CAL group had higher achievement and attitude scores, and they perceived their Prolog-based CAL classroom environments more favourably in terms of greater gender equity, investigation, innovation and resource adequacy.
BACKGROUND

Effect Sizes in Past Research on CAL

Although there has been considerable optimism about the potential of CAL to revolutionise education in terms of its potential to promote student achievement and attitudes, a myriad of evaluation studies of CAL have failed to support this optimism. Table 1 summarises the results of a set of 11 different past meta-analyses reviewed by Fraser, Walberg, Welch and Hattie [1]. These meta-analyses taken together encompassed 557 individual studies and tested 566 relationships involving the efficacy of CAL in terms of promoting student achievement. This table shows that the effect sizes for these meta-analyses, expressed as correlation coefficients, ranged from 0.03 to 0.26 for different meta-analyses and had a grand mean of 0.15.

The average effect size for these 11 meta-analyses, expressed in terms of the number of standard deviations of difference between a CAL and a control group, was found to be approximately 0.3. This value of 0.3 specifically for CAL studies is reasonably consistent with the synthesis [1] of 134 past meta-analyses of a wide range of educational variables which indicated an average effect size in education of approximately 0.4 standard deviations.

More recently, Roblyer, Castine and King's meta-analyses of CAL [3] revealed an average effect size of only 0.4 standard deviations, whereas Kulik and Kulik's [2] meta-analysis revealed a comparable value of 0.3 standard deviations.
Although past research has revealed only consistently modest effect size, it should be noted that many of the individual studies were conducted some years ago and the innovation in CAL often involved drill-and-practice. Consequently, there is a need to investigate the efficacy of more recent innovations in CAL.

Although past evaluative research on CAL has been plentiful, the investigation of effect sizes reported in this article is distinctive in three important ways. First, the innovation considered in this study involved micro-Prolog-based CAL, which has been the focus of very few past evaluations. Second, in addition to investigating the impact of CAL on traditional achievement and attitude outcomes, the present study appears to be the first which also evaluated CAL in terms of its impact on classroom psychosocial environment. Third, the study was conducted in the unique milieu of Singapore secondary schools.

The Singapore Context

In many ways, the Singapore education system is different from that of the United Kingdom and many other countries. For example, the Singapore system is highly centralised (with almost every school using the same prescribed textbooks), the mode of instruction is essentially expository and achievement-oriented, and schools are graded for excellence based solely on the students' academic performance. In such a competitive environment, there often is a quest for excellence and a thirst for knowledge and high performance. Educational applications of computers were introduced fairly extensively into the schools only in 1980. The focus initially was on teaching computer science as a subject at the pre-university level and low-level computer literacy courses at the secondary school level; there was no extensive computer education program at the primary level. CAL was introduced to the schools only in 1986, and currently schools use CAL in science, mathematics and English for remediation and enrichment. Singaporean schools do not develop their own CAL courseware.
In Singapore, there is a dearth of research on the effects of CAL. To date, only three research studies have been conducted. On the whole, positive effects of CAL have been found. However, because the samples in these studies were small and non-random, and because of the short duration of the experimental treatment and the narrow scope of the content selected, their findings should be interpreted with caution. Furthermore, the sample selected in one of the studies was from one of the best schools in the country, thus limiting the generalisability of the results obtained. Thus, research data on the use of CAL in Singapore is sparse, of questionable validity and of low generalisability. Because of the lack of dependable research information, and in view of the potential that research conducted in the unique educational context of Singapore has for our understanding of computer-assisted learning environments, the present study was conceptualised and conducted.

Classroom Environment

The field of classroom psychosocial environment now is established as a thriving area of research [4-5]. Nonetheless although there has been a revolution in recent years in terms of both the scope and variety of the roles which computer-assisted learning has played within education, the field of classroom learning environment [4-5] has lagged behind in two major ways. First, hitherto there has existed no learning environment instrument which has been tailor-made specifically for use in classrooms using computer-assisted learning. Second, innovations in computer-assisted learning rarely have been evaluated in terms of their impact on the nature of the classroom learning environment as perceived by students [6]. The present paper fills some of the gaps by reporting the evaluation of an innovation in computer-assisted learning in terms of its impact on classroom environment, in addition to reporting its impact on student achievement and attitudes.
PROLOG-BASED CAL ENVIRONMENTS

The innovation in computer-assisted learning which formed the focus of this study was distinctive in that it made use of micro-Prolog, the microcomputer version of Prolog, a high-level fifth-generation declarative language, which was developed for artificial intelligence research and which has been found to be robust, compact and accessible to students in classroom applications [7].

The University of Exeter School of Education has been actively involved in the use of micro-Prolog to help teach history, other humanities subjects and pupils with specific learning difficulties (dyslexics). The approach adopted by the project team has been to use micro-Prolog to implement application shells or authoring software that provide a programming environment in which teachers and students design and write 'programs' [8].

One of the Exeter programs, LINX, was used to create CAL modules for this study. LINX is a specific-purpose application shell which enables users to create textual programs with a branching tree structure. LINX is logical, powerful, well-structured and flexible [9]. It provides a structure for writing CAL simulations or other programs which involve the user in decision-making. Teachers can write CAL programs using LINX by inserting their own structure and data. LINX also allows students to explore decision-making by providing a framework into which they can enter a network of possible decisions and consequences. A LINX program takes the form of a decision tree with a branching structure of links and choices.

The CAL courseware which formed the focus of the present study covered the topic of decision-making in social science and was designed specifically for a six-week segment in the Singapore school social science syllabus. In order to
enhance the validity of conclusions from the evaluative study, a control group, which studied the same topic via the direct expository teaching methods common in Singapore, also was included in the study.

The CAL approach consisted of a sequence of instructions consisting of learning activities for the concept of decision-making, followed by a prescribed set of exercises. During the CAL lessons, students were presented with the necessary information. Students also were guided with instructions on the computer screens, were provided with practice exercises, and had their learning assessed. Feedback was provided contingent upon correct and incorrect responding. In other words, the program is user-friendly and easy to run, and the looping and branching options allow students to check, retrace or skip parts of the program. The student also can print a copy of the current screen, or the whole lesson. But whatever the students do, one of the function keys allows the students to look at the progress or various stages of their lessons, or the number of loops or branchings that they have performed. This can be displayed on the screen, or recorded on the diskette or printed out.

**EFFECT SIZES FOR ACHIEVEMENT AND ATTITUDES**

The study involved 12 teachers, each in a different randomly selected school. In order to reduce the 'teacher effect', each teacher taught one experimental and one control class. All schools were coeducational in order to allow an unconfounded test of whether CAL was differentially effective for boys and girls. The total number of students in these 24 classes was 671 (348 in the experimental group and 323 in the control group). Approximately equal number of males and females made up the sample. Students were slow learners in their second year of high school in Singapore.
The instruments used to measure achievement in and attitude towards social science were developed and validated for the present study. The Social Science Achievement Test (SSAT) is a 30-item multiple-choice test which assesses the concept of decision-making in the social sciences. The Semantic Differential Inventory (SDI) is a 20-item semantic differential instrument which measures students' attitudes towards learning social science. The alpha reliability coefficient for the whole sample was found to be 0.95 for the 30-item SSAT and 0.94 for the 20-item SDI.

To investigate differences in students' achievement and attitudes between the CAL and control group, ANCOVA procedures were computed separately for the achievement and attitude measures. The three covariates used were students' primary school leaving examination scores, their semestral achievement scores and their pretest scores on the achievement or attitude measure.

For achievement, the ANCOVA yielded a significant $F$ value of 5824.59 ($p<0.01$) which confirmed the presence of a reliable treatment effect. With the covariates partialled out of the total variance, the $\eta^2$ statistic (i.e., the amount of variance accounted for) had the high value of 0.90. That is, about nine-tenths of the variance in posttest achievement scores could be attributed uniquely to the treatment beyond that accounted for by the three covariates. The effect size (in terms of the difference between the means of the CAL and control groups divided by the standard deviation of the control group) had the very large value of 3.5, with the difference favouring the CAL group.

For the attitude measure, the ANCOVA yielded the $F$ value of 630.13 ($p<0.01$) and the $\eta^2$ statistic was 0.52 after covariates were partialled out. The effect size was 1.4 standard deviations and differences favoured the CAL group.
EFFECT SIZES FOR CLASSROOM ENVIRONMENT

Over the past 20 years or so, considerable progress has been achieved in the field of classroom environment research [5]. Studies in numerous countries have shown that students' perceptions of their classroom environment account for appreciable amounts of outcome variance, often beyond that attributable to student aptitude. Curriculum evaluators have found that student perceptions of classroom environment have differentiated revealingly between alternative curricula even when outcome measures have shown negligible differences [10].

While research on students' perceptions of psychosocial characteristics of their classrooms is not new, studying the impact of microcomputers on classroom learning environments is new. Most research on computers in school and classroom settings has been concerned with gender equity in computer learning [11], the effectiveness of computer-assisted learning in terms of cognitive and affective outcomes [2], and student preference for software. Little research has involved the impact of computers on learning environments in education.

Because hitherto no classroom environment instrument tailor made for CAL settings existed, a new instrument was developed for the purposes of the present research. The initial development of the new questionnaire took cognisance of Moos's three general dimensions as they apply to all human environments [12]. These three general dimensions are Relationship Dimensions (the nature and intensity of personal relationships within the environment), Personal Development Dimensions (the basic directions along which personal growth and self-enhancement tend to occur) and System Maintenance and System Change Dimensions (the extent to which the environment is orderly, clear in expectations, maintains control and is responsive to change). Table 2 shows that the four scales in the final version of the instrument provide reasonable coverage of the three different basic types of
dimensions proposed by Moos. The development of the instrument also benefitted from a review of the literature for the purpose of identifying scales that are considered important in the unique environment of computer-assisted learning, and from comments and feedback obtained from social science teachers and second year students at high school in Singapore on draft versions of sets of items.

To investigate differences in students' perceptions of classroom environment between the experimental (computer) and control (non-computer) group, ANCOVA procedures were used again for the same sample with the same three covariates. A separate ANCOVA was computed for each of the four scales. A summary of the results for significance of differences between the perceptions of the computer and non-computer groups are reported in Table 3. The effect size (i.e., the number of standard deviations of difference between the experimental CAL group and the control group) for each of the classroom environment scales also is reported in Table 3.

Significant differences \((p<0.01)\) emerged in the students' perceptions between the computer and non-computer group for all of the four scales, namely, Gender Equity, Investigation, Innovation and Resource Adequacy. Table 3 also
reveals that the use of Prolog-based computer-assisted learning was associated with an effect size of 1.0 standard deviations for Gender Equity, 1.9 standard deviations for Investigation, 1.7 standard deviations for Innovation and 1.5 standard deviations for Resource Adequacy. These differences favoured the CAL group in every case. Relative to control classes, the CAL classes were perceived to have greater gender equity, investigation, innovation and resource adequacy. The present study's findings in the context of computer-assisted learning are consistent with past studies on non-CAL classroom environments [4, 10] in which learning environment measures have proved useful in curriculum evaluation.

DISCUSSION

This study of the use of computer-assisted learning in school social science education in Singapore is distinctive because, first, the courseware made use of micro-Prolog, second, research on CAL in social science education in Singapore hitherto has been non-existent and, third, past evaluations of CAL have not included an investigation of the effects of using CAL on the classroom psychosocial environment. It was found that appropriately-designed Prolog-based CAL can be an effective instructional method in the classroom milieu and that the educational application of micro-Prolog as a learning tool in social science classrooms can be efficacious. While a myriad of studies have examined the effects of CAL, the number of studies examining the effect of Prolog-based CAL is meager at best. Hopefully, this study will serve as a catalyst for further research into the use of Prolog in the classroom.

No systematic attempts previously have been made to examine the effects of CAL on learning environment characteristics. Little has been studied or is known about the impact of microcomputers on learning climates in education. This study
responds to the plea by MacGregor [6] and Lancy [14] to fill the lacunae represented by the study of computers and their impact on students. This study is significant because, in contrast to previous research, it uses Prolog-based courseware developed by the investigators in assessing computer-assisted learning environments in schools. One of the study's original contributions was that a new classroom environment instrument was developed and validated specifically for the unique setting of computer-assisted learning. The scales in this instrument were found to display adequate factorial validity, internal consistency reliability, discriminant validity and predictive validity (in terms of being significantly related to student outcome scores). Also each scale differentiated significantly between the perceptions of students in different classrooms. It is likely that other researchers will find this new instrument useful in future studies of CAL classroom environments.

One of the major findings was that, in contrast to past research, the use of CAL in this study led to massive impact in terms of achievement (effect size of 3.5 standard deviations), attitudes (1.4 standard deviations) and classroom environment (ranging from 1.0 to 1.9 standard deviations). These figures can be compared with the effect sizes of 0.3-0.4 found in past meta-analyses of the effect of CAL on student achievement [1-3]. The large effect sizes arising from this study could be attributed to the way in which the Prolog-based CAL courseware took cognisance of the curriculum objectives of the schools, the prerequisites of following the exact curriculum topics, and the integration of the courseware with the syllabus requirements of the schools. It seems that appropriate computer-based teaching is effective with low aptitude students, which is consistent with the meta-analysis of CAL effectiveness of Bangert-Drowns, Kulik and Kulik [13]. Replication and further related research is recommended, however, because of the uniqueness of the Singapore milieu.
References


<table>
<thead>
<tr>
<th>Area covered</th>
<th>No. of studies</th>
<th>No. of relationships</th>
<th>Overall correlation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>College level</td>
<td>59</td>
<td>54</td>
<td>0.12</td>
<td>Kulik et al. [1980]</td>
</tr>
<tr>
<td>High school</td>
<td>51</td>
<td>48</td>
<td>0.16</td>
<td>Kulik et al. [1983]</td>
</tr>
<tr>
<td>Mathematics</td>
<td>153</td>
<td>89</td>
<td>0.20</td>
<td>Hartley [1980]</td>
</tr>
<tr>
<td>Mathematics</td>
<td>15</td>
<td>126</td>
<td>0.03</td>
<td>Leong [1980]</td>
</tr>
<tr>
<td>Science</td>
<td>11</td>
<td>14</td>
<td>0.10</td>
<td>Aiello &amp; Wolfe [1980]</td>
</tr>
<tr>
<td>Science</td>
<td>14</td>
<td>14</td>
<td>0.06</td>
<td>Willett et al. [1983]</td>
</tr>
<tr>
<td>Mathematics &amp; science</td>
<td>101</td>
<td>68</td>
<td>0.26</td>
<td>Curbelo [1984]</td>
</tr>
<tr>
<td>Elementary schools</td>
<td>48</td>
<td>48</td>
<td>0.15</td>
<td>Niemiec [1985]</td>
</tr>
<tr>
<td>Elementary schools</td>
<td>25</td>
<td>25</td>
<td>0.23</td>
<td>Kulik et al. [1984]</td>
</tr>
<tr>
<td>Drill &amp; practice</td>
<td>40</td>
<td>40</td>
<td>0.17</td>
<td>Burns &amp; Bozeman [1981]</td>
</tr>
<tr>
<td>Tutorial work</td>
<td>40</td>
<td>40</td>
<td>0.22</td>
<td>Burns &amp; Bozeman [1981]</td>
</tr>
<tr>
<td>Overall</td>
<td>557</td>
<td>566</td>
<td>0.15</td>
<td></td>
</tr>
</tbody>
</table>

Based on Fraser et al. [1]
Table 2: Descriptive information for classroom environment scales

<table>
<thead>
<tr>
<th>Scale Name</th>
<th>Description</th>
<th>Moo's classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender Equity</td>
<td>Extent to which boys and girls are treated equally by the teacher</td>
<td>Relationship</td>
</tr>
<tr>
<td>Investigation</td>
<td>Extent to which the skills and processes of inquiry are used in problem-solving and investigation</td>
<td>Personal Development</td>
</tr>
<tr>
<td>Innovation</td>
<td>Extent to which the teacher plans new and varying activities and techniques, and encourages students to think creatively</td>
<td>System Maintenance</td>
</tr>
<tr>
<td>Resource Adequacy</td>
<td>Extent to which the computer hardware and software are adequate</td>
<td>System Maintenance</td>
</tr>
</tbody>
</table>

Table 3: A comparison of experimental and control groups on classroom environment scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>No. of items</th>
<th>$F$</th>
<th>Effect size$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender Equity</td>
<td>8</td>
<td>219.66**</td>
<td>1.0</td>
</tr>
<tr>
<td>Investigation</td>
<td>8</td>
<td>809.14**</td>
<td>1.9</td>
</tr>
<tr>
<td>Innovation</td>
<td>8</td>
<td>703.63**</td>
<td>1.7</td>
</tr>
<tr>
<td>Resource Adequacy</td>
<td>8</td>
<td>504.28**</td>
<td>1.5</td>
</tr>
</tbody>
</table>

$** p<0.01$

$^a$ The effect size is the difference between the means of the experimental and control groups divided by the standard deviation of the control group.